

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 07:13:21 ; Search time 1916 Seconds

(without alignments)  
13078.035 Million cell updates/sec

Title: US-09-934-249-3

Sequence: 1 atgcaccgccttgatggtggtc.....aacagaagagcaccccttc 861

Scoring table: OLIGO-MNC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*  
29: em.vi:\*  
30: em.htg.hum:\*  
31: em.htg.in:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.rtd:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.syl:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

Score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	861	100.0	861	6	AX392419	AX392419 Sequence
2	861	100.0	1321	6	AX392417	AX392417 Sequence
3	861	100.0	4839	9	AF305616	AF305616 Homo sapi
4	790	91.8	1061	9	BC015918	BC015918 Homo sapi
5	752	87.3	1141	9	AF224278	AF224278 Homo sapi
6	752	87.3	1818	9	AY128643	AY128643 Homo sapi
7	545	63.3	61505	9	AF305426	AF305426 Homo sapi
8	545	63.3	130435	9	HS71807	HS71807 Homo sapi
9	401	46.6	408	6	AX071267	AX071267 Sequence
10	302	35.1	693	6	AX392430	AX392430 Sequence
11	271	31.5	812	6	AX011709	AX011709 Sequence
12	156	18.1	150224	9	HSJ105917	HSJ105917 Sequence
13	51	5.9	51	6	AX199565	AX199565 Sequence
14	44	5.1	651	10	AF220208	AF220208 Mus muscu
15	44	5.1	878	6	AX392428	AX392428 Sequence
16	44	5.1	156075	2	AL837509	AL837509 Mus muscu
17	44	5.1	167758	2	AC110189	AC110189 Mus muscu
18	41	4.8	475	6	AX392431	AX392431 Sequence
19	34	3.9	159824	2	AC111878	AC111878 Rattus no
20	32	3.7	756	10	AB045588S2	AB045588 Rattus no
21	32	3.7	56475	9	AL353658	AL353658 Human DNA
22	32	3.7	60989	2	AC094865	AC094865 Rattus no
23	32	3.7	111961	2	AC126194	AC126194 Rattus no
24	32	3.7	163085	2	AC106160	AC106160 Rattus no
25	32	3.7	165438	2	AC126988	AC126988 Rattus no
26	32	3.7	172525	2	AC008514	AC008514 Homo sapi
27	32	3.7	179372	2	AC040922	AC040922 Homo sapi
28	32	3.7	198582	2	AC005291	AC005291 Homo sapi
29	32	3.7	223029	2	AC111410	AC111410 Rattus no
30	31	3.6	67987	10	AL672157	AL672157 Mouse DNA
31	31	3.6	70811	2	AC117542	AC117542 Mus muscu
32	31	3.6	96240	2	AC110445	AC110445 Rattus no
33	31	3.6	114287	2	AP000590	AP000590 Homo sapi
34	31	3.6	162812	2	AC129076	AC129076 Mus muscu
35	31	3.6	166474	2	AC107431	AC107431 Rattus no
36	31	3.6	169163	2	AL606807	AL606807 Homo sapi
37	31	3.6	183301	2	AC092487	AC092487 Homo sapi
38	31	3.6	189982	2	AC079023	AC079023 Mus muscu
39	31	3.6	210608	2	AC006028	AC006028 Homo sapi
40	30	3.5	79951	9	HSJ908M14	HSJ908M14 Human DNA
41	30	3.5	85546	2	AC111932	AC111932 Rattus no
42	30	3.5	99497	9	HS29316	HS29316 Human DNA
43	30	3.5	145241	2	AC117110	AC117110 Rattus no
44	30	3.5	156013	2	AC126882	AC126882 Rattus no
45	30	3.5	167936	2	AC106143	AC106143 Rattus no

#### ALIGNMENTS

RESULT 1	AX392419	861 bp	DNA	Linear	PAT 23-MAR-2002
LOCUS	AX392419				
DEFINITION	Sequence 3 from Patent WO0216416.				
ACCESSION	AX392419				
VERSION	AX392419.1	GI:19700734			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and Turi, T.G.				
	Diagnosis and treatment of cardiovascular conditions				

Pred. No. is the number of results predicted by chance to have a





OY 841 AACAGAAAGACACCTCTC 861  
 Db 1161 AACAGAAAGACACCTCTC 1181

RESULT 4  
 LOCUS BC015918 1061 bp mRNA linear PRI 17-OCT-2001  
 DEFINITION Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.  
 ACCESSION BC015918  
 VERSION BC015918.1 GI:16198474  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1061)  
 Strausberg, R.  
 Direct Submission  
 Submitted (15-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapsd-remail.nih.gov](mailto:cgapsd-remail.nih.gov)  
 Tissue Procurement: DCTD/DRP  
 DNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bgsc.bc.ca  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven  
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedi, Jacqueline  
 Schein, Duane Smalrus, Michael Smith, Lorraine Spence, Jeff Stolt,  
 Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES  
 Source  
 Location/Qualifiers  
 1..1061  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:20374 IMAGE:4559576"  
 /rissue\_type="Kidney, renal cell adenocarcinoma"  
 /clone\_lib="NIH\_MGC\_14"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 131..844  
 /codon\_start=1  
 /product="unknown (protein for MGC:20374)"  
 /protein\_id="AAH15918.1"  
 /db\_xref="GI:16198475"  
 /translation="MMVAVVITCLSHYKLSARFISRHSGRRREDAISSGCLMP  
 SESTVSGNGIPEPOVYAPRPRTDLAVPRAQREPRFRPTPTVLOHETDLPPTLS  
 SDGEPPYOGPCAIORPEOOLELNREVRAPPNRTITDSULMSARLGGCPSS  
 NSGISATCVSGSRMBEPPPTTISEVIGHPGSSFOHQSSGPPSLLEGTRLHTHTIAP  
 LESAHTSKRKDKQKGPPL"  
 BASE COUNT 239 a 350 c 311 g 161 t

Query Match 91.8%; Score 790; DB 9; Length 1061;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 840; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 21 CAACAGCACCAGCCGCGCCGCGGAGGCCAATGTCCTCTGACCTGCATCTGCA 80  
 Db 1 CAACAGCACCAGCCGCGCCGCGGAGGCCAATGTCCTCTGACCTGCATCTGCA 60  
 OY 81 ACCGCTCTTTGTCACAGCATGGAGATCAACGAGCTGGAGTTGTTCAATCATCATCAT 140  
 Db 61 ACCGCTCTTTGTCACAGCATGGAGATCAACGAGCTGGAGTTGTTCAATCATCATCAT 120  
 OY 141 CGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 200  
 Db 121 CGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 OY 201 GTCTGCAGGTCCTTCATCAGCCGACAGCCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 260  
 Db 181 GTCTGCAGGTCCTTCATCAGCCGACAGCCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 OY 261 CTCAGAGAGATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320  
 Db 241 CTCAGAGAGATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 OY 321 GCAGGTCTAGCCGCGCCGCTGCGCCACAGCAGCCGCTGCGCCGCTGCGCCGCTGCGCCG 380  
 Db 301 GCAGGTCTAGCCGCGCCGCTGCGCCACAGCAGCCGCTGCGCCGCTGCGCCGCTGCGCCG 360  
 OY 381 GGAGGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCC 440  
 Db 361 GGAGGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCC 420  
 OY 441 ACCCAGCATCTGCTGTGATGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500  
 Db 421 GCCCAGCATCTGCTGTGATGAG 480  
 OY 501 CCAGCTTGGGAGCCCGGAGCAGCAGCTGGAACTGAACCGGAGAGTGGTGGCGCAGCCGCC 560  
 Db 481 CCAGCTTGGGAGCCCGGAGCAGCAGCTGGAACTGAACCGGAGAGTGGTGGCGCAGCCGCC 540  
 OY 561 AAACGAGACATCTTCCAGCACTGACAGCTGATGATGATGATGATGATGATGATGATGATGAT 620  
 Db 541 AAACGAGACATCTTCCAGCACTGACAGCTGATGATGATGATGATGATGATGATGATGATGAT 600  
 OY 621 CCCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680  
 Db 601 CCCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 OY 681 GCGCGCGCCACCTTACAGCGAGAGTCAATGCGCCACTACCGGGGTCTCTCTTCCAGCAGCA 740  
 Db 661 GCGCGCGCCACCTTACAGCGAGAGTCAATGCGCCACTACCGGGGTCTCTCTTCCAGCAGCA 720  
 OY 741 GCAGAGCATGTGGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 800  
 Db 721 GCAGAGCATGTGGCGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 OY 801 GCCCTTGAAGAGCGCAGCAGCTGTGAGCAAGAGAGATTAACAGAAAGAGACACCTCTCT 860  
 Db 781 GCCCTTGAAGAGCGCAGCAGCTGTGAGCAAGAGAGATTAACAGAAAGAGACACCTCTCT 840  
 OY 861 C 861  
 Db 841 C 841

RESULT 5  
 AF224278 1141 bp mRNA linear PRI 18-JUL-2000  
 LOCUS Homo sapiens PMERAI protein (PMERAI) mRNA, complete cds.  
 DEFINITION AF224278  
 ACCESSION AF224278.1 GI:9255808  
 VERSION AF224278.1 GI:9255808  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



```

REFERENCE      1 (bases 1 to 1141)
AUTHORS        Xu,L., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
               Moul,J.W. and Srivastava,S.
TITLE          A novel androgen-regulated gene, PMEPA1, located on chromosome
               20q13 exhibits high level expression in prostate
JOURNAL        Genomics 66 (3), 257-263 (2000)
MEDLINE        20334621
PUBMED        10873380
REFERENCE      2 (bases 1 to 1141)
AUTHORS        Xu,L., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
               Moul,J.W. and Srivastava,S.
TITLE          Direct Submission
JOURNAL        Submitted (12-JAN-2000) CPDR, USUHS, 1530 East Jefferson Street,
               Rockville, MD 20852, USA
FEATURES       source
               location/Qualifiers
               1..1141
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="20"
               /map="20q13.31-13.33"
               /cell_line="LNCaP"
               1..1141
               /gene="PMEPA1"
               /note="type 1b transmembrane protein; expression is
               induced in response to the synthetic androgen, R1881;
               expression is abundant in, and restricted to, prostate
               glandular epithelial cells; similar to the predicted
               protein encoded by sequence deposited at GenBank Accession
               Number NP_004329"
               /codon_start=1
               /product="PMEPA1 protein"
               /protein_id="AA86322.1"
               /db_xref="GI:9255809"
               /translation="MALEFVQIIIVVVMVVVITCLSHYKLSARSFISRHSG
               RREDLSSBGLPSESVNGIPPEPVYAPRPDRLAAPPFQOREFHFPQY
               PYLQHEIDLPPTISLDGEEPPYQCGTOLADPQOLENESVAPNPRTIFSD
               LMSARLGGPCPPSSNSGTSATCYSGGRMEGPPPTSEVIGHYPPSSFHQOSSGPP
               SLLEGRHLHTHTAPLESAAIWSKEXDKQGHPL"
BASE COUNT    271 a 350 c 336 g 184 t
ORIGIN
Query Match   87.3%: Score 752; DB 9; Length 1141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 110 CGAGCTGAGATTGTTGATCATCATCATCGTGTGATGATGATGATGATG 169
    |||||||
Db 100 CGAGCTGAGATTGTTGATCATCATCATCGTGTGATGATGATGATGATG 159
QY 170 TGATCAGTGCCTGCTGAGCAGTACAAAGCTGTGTGACAGGCTCTCATCAGCGGACACA 229
    |||||||
Db 160 TGATCAGTGCCTGCTGAGCAGTACAAAGCTGTGTGACAGGCTCTCATCAGCGGACACA 219
QY 230 GCCAGGGGGGAGAGAGAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTTGAGAGA 289
    |||||||
Db 220 GCCAGGGGGGAGAGAGAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTTGAGAGA 279
QY 290 GCACAGTGTCAAGGACGAGATCCAGAGCGGAGGTGTACGCCCGGCTGGGCCACAG 349
    |||||||
Db 280 GCACAGTGTCAAGGACGAGATCCAGAGCGGAGGTGTACGCCCGGCTGGGCCACAG 339
QY 350 ACCGCTGGCCGTCGCCCTTCGCCAGCGGAGGAGCGCTTCACCGCTTCACGCCACCT 409
    |||||||
Db 340 ACCGCTGGCCGTCGCCCTTCGCCAGCGGAGGAGCGCTTCACCGCTTCACGCCACCT 399
QY 410 ATCCGTACTCTGACAGAGATCGACCTGCGCACCCACCATCTCGCTGTCAACAGGGGAGG 469
    |||||||
Db 400 ATCCGTACTCTGACAGAGATCGACCTGCGCACCCACCATCTCGCTGTCAACAGGGGAGG 459
QY 470 AGCCCCACCTTACAGAGGGCCCTGACACCTTCAGACTTCGGAGACCCGAGCAGCTGG 529
    |||||||

```

```

Db 460 AGCCCCACCTTACAGAGGGCCCTGACACCTTCAGACTTCGGAGACCCGAGCAGCTGG 519
QY 530 AACTGAAACCGGAGTGTGCGGCGACCCCAACACAGACCATCTTGACAGTACCTGA 589
    |||||||
Db 520 AACTGAAACCGGAGTGTGCGGCGACCCCAACACAGACCATCTTGACAGTACCTGA 579
QY 590 TGTATAGTGCAGAGCTGGGCGGCGGCTGCCCCCAGCAGTAACTCGGGCATCAGCGCA 649
    |||||||
Db 580 TGTATAGTGCAGAGCTGGGCGGCGGCTGCCCCCAGCAGTAACTCGGGCATCAGCGCA 639
QY 650 CTGTCTACGGCAGCGGCGGCGCATGTAGAGGGCGCGCCACCTACAGCAGGTCATCG 709
    |||||||
Db 640 CTGTCTACGGCAGCGGCGGCGCATGTAGAGGGCGCGCCACCTACAGCAGGTCATCG 699
QY 710 GCCACTACCCGGGCTCCCTTCAGACACGACAGAGAGAGGGCGGCTTCCTTGGCTGG 769
    |||||||
Db 700 GCCACTACCCGGGCTCCCTTCAGACACGACAGAGAGAGGGCGGCTTCCTTGGCTGG 759
QY 770 AGGGACCCGCGCTCCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGAGACA 829
    |||||||
Db 760 AGGGACCCGCGCTCCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGAGACA 819
QY 830 AAGAGAAAGATAAACAGAAAGACACCTCTC 861
    |||||||
Db 820 AAGAGAAAGATAAACAGAAAGACACCTCTC 851
RESULT 6
LOCUS      AY128643 1818 bp mRNA linear PRI 05-AUG-2002
DEFINITION Homo sapiens PMEPA1 variant A protein mRNA, complete cds.
ACCESSION  AY128643
VERSION     AY128643.1 GI:22121998
KEYWORDS   human.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 1818)
            Brunschwig,E.B., Wilson,K., Mack,D., Dawson,D., Lawrence,E.,
            Willison,J.K.V., Lu,S., Nosrati,A., Swinder,S., Beard,L.,
            Lutterbaugh,J.D., Willis,J., Platzer,P. and Markowitz,S.
            Direct Submission
            Submitted (03-JUL-2002) Department of Medicine, Case Western
            Reserve University/Howard Hughes Medical Institute, 11001 Cedar
            Ave., Cleveland, OH 44106, USA
FEATURES   source
            location/Qualifiers
            1..1818
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20q13.31-13.33"
            145..858
            /codon_start=1
            /product="PMEPA1 variant A protein"
            /protein_id="AA86927.1"
            /db_xref="GI:22121998"
            /translation="MWVVVVITCLSHYKLSARSFISRHSGRRREDLSSBGLP
            SESYVSGNGIPPEPVYAPRPDRLAAPPFQOREFHFPQYLYQHEIDLPPTISL
            SDGEEPPYQCGTOLADPQOLENESVAPNPRTIFSDLMSARLGGPCPPSS
            NSGTSATCYSGGRMEGPPPTSEVIGHYPPSSFHQOSSGPPSLLEGRHLHTHTAP
            LESAAIWSKEXDKQGHPL"
BASE COUNT 459 a 467 c 494 g 398 t
ORIGIN
Query Match   87.3%: Score 752; DB 9; Length 1818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 110 CGAGCTGAGATTGTTGATCATCATCATCGTGTGATGATGATGATGATG 169
    |||||||
Db 104 CGAGCTGAGATTGTTGATCATCATCATCGTGTGATGATGATGATGATG 163

```

QY	170	TGATCATGCGCCGCTGCGAGCCTATACAGCTGTCTGACGGTCCCTTCATCAGCCGGCACA	229
Db	164	TGATCATGCGCCGCTGCGAGCCTATACAGCTGTCTGACGGTCCCTTCATCAGCCGGCACA	223
QY	230	GCCAGGGGGGAGAGAGAGAAGATGCCCTGTCTCACAAGGATGCCCTGTGGCCCTCGGAGA	289
Db	224	GCCAGGGGGGAGAGAGAGAAGATGCCCTGTCTCACAAGGATGCCCTGTGGCCCTCGGAGA	283
QY	290	GCACAGTGTACAGGCAACGGGAATCCCAAGACCCGAGGTCTACGCGCCCGCTCGGCCACCG	349
Db	284	GCACAGTGTACAGGCAACGGGAATCCCAAGACCCGAGGTCTACGCGCCCGCTCGGCCACCG	343
QY	350	ACCGCGCTGGCGCTGCGCCCTTTCGCCACGCGGAGCGCTTCCACCGCTTCCAGCCCACTT	409
Db	344	ACCGCGCTGGCGCTGCGCCCTTTCGCCACGCGGAGCGCTTCCACCGCTTCCAGCCCACTT	403
QY	410	ATCCGCTACTGCAAGCAGCAGATCGACTTCCACCAACCATCTTCGTGTGCAGACGGGGAGG	469
Db	404	ATCCGCTACTGCAAGCAGCAGATCGACTTCCACCAACCATCTTCGTGTGCAGACGGGGAGG	463
QY	470	AGCCCCCACTTACAGAGGGGCCCTGCACCTCCAGCTTCGGGGACCCCGAGACGACACTGG	529
Db	464	AGCCCCCACTTACAGAGGGGCCCTGCACCTCCAGCTTCGGGGACCCCGAGACGACACTGG	523
QY	530	AACCTGAACCGGAGATCGGTGCGCGCACCCCCAAACAGAACCATCTTGCACAGTACACTTA	589
Db	524	AACCTGAACCGGAGATCGGTGCGCGCACCCCCAAACAGAACCATCTTGCACAGTACACTTA	583
QY	590	TGGATATGTCACGAGCGGGGCGGCCCTTGCGCCCCCGACGATACGTCGGGCAATCAGGCGCA	649
Db	584	TGGATATGTCACGAGCGGGGCGGCCCTTGCGCCCCCGACGATACGTCGGGCAATCAGGCGCA	643
QY	650	CCTGCTACGGGACGCGGGGGCGCATATGAGAGGGGCGCGGCCCACTACAGCAGAGTATATCG	709
Db	644	CCTGCTACGGGACGCGGGGGCGCATATGAGAGGGGCGCGGCCCACTACAGCAGAGTATATCG	703
QY	710	GCCACTACCCGGGGGTCTCTCTTCACAGCACACAGCAGAGTGGGCCCGCTCTCTGTCTGG	769
Db	704	GCCACTACCCGGGGGTCTCTCTTCACAGCACACAGCAGAGTGGGCCCGCTCTCTGTCTGG	763
QY	770	AGGGGACCCGGGCTGCACACACACACATCGCGCCCTTAGAGAGCGCAGGCATCTGGAGCA	829
Db	764	AGGGGACCCGGGCTGCACACACACACATCGCGCCCTTAGAGAGCGCAGGCATCTGGAGCA	823
QY	830	AAGAGAAGGATTAACAGAAAGACACACCCCTCTC	861
Db	824	AAGAGAAGGATTAACAGAAAGACACACCCCTCTC	855
RESULT 7			
LOCUS	AF305426		
DEFINITION	Homo sapiens solid tumor-associated 1 protein (STAG1/PMEP1) gene,		
ACCESSION	AF305426		
VERSION	AF305426.1		
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 61505)		
DATE	Ree,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.		
COMMENT	Identification and characterization of a novel gene, STAG1,		
	up-regulated in renal cell carcinoma and other solid tumours		
	Unpublished		
	2 (bases 1 to 61505)		
	Ree,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.		
	Direct Submission		
	Submitted (13-SEP-2000) Centre for Molecular Biotechnology,		
	Queensland University of Technology, 2 George St, Brisbane, QLD		
	4001, Australia		
	Location/Qualifiers		
FEATURES			

source	gene	mrna	CDS	BASE COUNT
1. .61505				13329
/organism="Homo sapiens"				a
/db_xref="taxon:9606"				16074
/chromosome="20"				c
/map="20q13.2-q13.33"				17430
1. .61505				g
/gene="STAG1/PMEPA1"				14672
join(1. .429,50206. .50360,56817. .56870,57305. .61505)				t
/gene="STAG1/PMEPA1"				
/product="solid tumor-associated 1 protein"				
join(321. .429,50206. .50360,56817. .56870,57305. .57850)				
/gene="STAG1/PMEPA1"				
/codon_start=1				
/product="solid tumor-associated 1 protein"				
/protein_id="AA109357.1"				
/db_xref="GI:15824469"				
/translation="MHRLMGVNSTAAANAAGPNVSCNCKRSLEFSQSMETLEFEVQI IIIVVAVMAYVVITCLLSHKLRARSFISRHSGRRREDALESGLCWPESEVSGG GIPPEVAVMPRPRIQLAVPFAQREHRRFPQTPYPLQHEIDLPRTSLDGEPP YOGCTCLRDPEOOLNEHRSVRAVPMRRTIFSDLDMSARLGGCCPSSSGISATQ YGSGRMEGAPRPMTSEVYIGHYFGSSFFQGGSSGPPSLLEGRLLMHTHTAPLESAAIWS KEREKOKGHP"				

Query Match	Similarity	63.3%	Score 545	DB 9	Length 61505
Best Local	Similarity	100.0%	Pred. NO. 4.2e-275		
Matches	545	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY 317	AGCCGAGGTCTACGCCCCGCTTCGGCCACGACGCGCTGGCGCTGCGCCCTTCGCCC				376
Db 57303	AGCCGAGGTCTACGCCCCGCTTCGGCCACGACGCGCTGGCGCTGCGCCCTTCGCCC				57362
QY 377	AGCGGGAGCGCTTCACCGCTTCAGGCCACTATCCGTAAGTGTGACACGAGATCGACC				436
Db 57363	AGCGGGAGCGCTTCACCGCTTCAGGCCACTATCCGTAAGTGTGACACGAGATCGACC				57422
QY 437	TGCGACCCACCATCTCGCTGTGACAGGGGAGAGACCCGACCTTACCGAGGACCCCTGCA				496
Db 57423	TGCGACCCACCATCTCGCTGTGACAGGGGAGAGACCCGACCTTACCGAGGACCCCTGCA				57482
QY 497	CCCTTCAGCTTGGGAGCCCGCGAGCAGCAGCTGGAACTGAAACCGGGAGTGGTGGCGCAC				556
Db 57483	CCCTTCAGCTTGGGAGCCCGCGAGCAGCAGCTGGAACTGAAACCGGGAGTGGTGGCGCAC				57542
QY 557	CCCCAAGAGAACCATCTTCGACAGTGAACCTGATGATAGTGGCCAGGGTGGGCGGCCCT				616
Db 57543	CCCCAAGAGAACCATCTTCGACAGTGAACCTGATGATAGTGGCCAGGGTGGGCGGCCCT				57602
QY 617	GCCCCCGCCAGCAGTAACTTCGGGGCATCACGCGCACCTGTCTAGCGGCAAGCGCGGGCATGG				676
Db 57603	GCCCCCGCCAGCAGTAACTTCGGGGCATCACGCGCACCTGTCTAGCGGCAAGCGCGGGCATGG				57662
QY 677	AGGGGCGCGCCGCCACCTACAGCAGAGGTATCGGCGCACTACCCGGGGTCTCTTCCAGC				736
Db 57663	AGGGGCGCGCCGCCACCTACAGCAGAGGTATCGGCGCACTACCCGGGGTCTCTTCCAGC				57722
QY 737	ACCGAGCAGAGAGTGGGGCGCGCCCTTCGCTGTGAGAGGGACCCGGGTCCACCAACACACA				796
Db 57723	ACCGAGCAGAGAGTGGGGCGCGCCCTTCGCTGTGAGAGGGACCCGGGTCCACCAACACACA				57782
QY 797	TGCGGCGCCCTAGAGAGCGCAGCAGCATCTTGAGACAAAGAAAGATTAACAGAAAGACACC				856
Db 57783	TGCGGCGCCCTAGAGAGCGCAGCAGCATCTTGAGACAAAGAAAGATTAACAGAAAGACACC				57842
QY 857	CTCTC 861				
Db 57843	CTCTC 57847				

DEFINITION Human DNA sequence from clone RP4-718J7 on chromosome 20q13.31-13.33 contains the PCR1 gene for soluble phosphoenolpyruvate carboxykinase 1, part of a novel gene similar to mouse DLM-1 (tumour stroma and activated macrophage protein), the 3' end of the TMEPAI gene encoding an androgen induced 1b transmembrane protein (PMEPAI), two putative novel genes, a Cpg island, ESTs, STSs and GSSs, complete sequence.

ACCESSION AL035541

VERSION AL035541.15 GI:11546043

KEYWORDS HTG; Cpg island; DLM-1; macrophage protein; PCR1; phosphoenolpyruvate carboxykinase; PMEPAI; TMEPAI; transmembrane protein.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 130435)

AUTHORS Sehra, H.

JOURNAL Direct Submission

COMMENT Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

On Dec 5, 2000 this sequence version replaced gi:10198628. requests: clonerequest@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/c-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP4-718J7 The true left end of clone RP5-1007E6 is at 71437 in this sequence. The true right end of clone RP4-579F20 is at 43945 in this sequence. RP4-718J7 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES

source

1..130435

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="20"

/map="q13.31-13.33"

/clone="RP4-718J7"

/clone\_1ib="RPCI-4"

319..478

repeat\_region

/note="5 copies 32 mer 67% conserved"

370..463

repeat\_region

/note="2 copies 47 mer 87% conserved"

427..626

repeat\_region

/note="4 copies 50 mer 71% conserved"

707..812

repeat\_region

/note="53 copies 2 mer cc 61% conserved"

967..2752

repeat\_region

/note="893 copies 2 mer gg 54% conserved"

982..2730

repeat\_region

/note="33 copies 53 mer 54% conserved"

1177..1820

repeat\_region

/note="4 copies 161 mer 64% conserved"

1204..1371

repeat\_region

/note="3 copies 56 mer 75% conserved"

1358..2757

repeat\_region

/note="28 copies 50 mer 54% conserved"

1675..2718

repeat\_region

/note="18 copies 58 mer 55% conserved"

1843..2346

repeat\_region

/note="9 copies 56 mer 64% conserved"

1852..2079

repeat\_region

/note="4 copies 57 mer 86% conserved"

2028..2632

repeat\_region

/note="11 copies 55 mer 60% conserved"

2087..2257

repeat\_region

/note="3 copies 57 mer 83% conserved"

2228..2497

repeat\_region

/note="5 copies 54 mer 75% conserved"

2583..2716

repeat\_region

/note="2 copies 67 mer 82% conserved"

3378..3426

repeat\_region

/note="LIMB4 repeat: matches 6088..6136 of consensus"

3799..3896

repeat\_region

/note="Charliel repeat: matches 681..781 of consensus"

5331..5793

repeat\_region

/note="MRT1B repeat: matches 14..466 of consensus"

5797..5988

repeat\_region

/note="6 copies 32 mer 86% conserved"

5855..5978

repeat\_region

/note="31 copies 4 mer gcac 61% conserved"

6474..6591

repeat\_region

/note="MIR repeat: matches 91..218 of consensus"

6592..6723

repeat\_region

/note="FLAK\_C repeat: matches 1..132 of consensus"

6952..7021

repeat\_region

/note="L2 repeat: matches 2637..2705 of consensus"

7358..7671

repeat\_region

/note="Alusx repeat: matches 1..312 of consensus"

8521..8554

repeat\_region

/note="Alu repeat: matches 1..34 of consensus"

8863..9217

repeat\_region

/note="MUT1D repeat: matches 105..505 of consensus"

9228..9746

repeat\_region

/note="L2 repeat: matches 1459..2028 of consensus"

9880..10089

repeat\_region

/note="L2 repeat: matches 2292..2511 of consensus"

10094..10206

repeat\_region

/note="Aluub repeat: matches 188..300 of consensus"

10263..10573

repeat\_region

/note="Aluub repeat: matches 1..308 of consensus"

10658..10753

repeat\_region

/note="L2 repeat: matches 2647..2749 of consensus"

11816..12380

repeat\_region

/note="LRR19B repeat: matches 1..580 of consensus"

12519..12813

repeat\_region

/note="LIMD2 repeat: matches 6032..6331 of consensus"

13559..13726

repeat\_region

/note="3 copies 56 mer 73% conserved"

13585..13716

repeat\_region

/note="22 copies 6 mer ctctct 67% conserved"

13586..13717

repeat\_region

/note="66 copies 2 mer tc 68% conserved"

13588..13715

repeat\_region

/note="4 copies 32 mer 78% conserved"

13589..13724

repeat\_region

/note="34 copies 4 mer cttt 77% conserved"

complement(13622..14142)

/note="match: GSS: Em:A0592603"

14313..14437

repeat\_region

/note="LRR16C repeat: matches 257..386 of consensus"

15015..15110

repeat\_region

/note="3 copies 32 mer 79% conserved"

15021..15110

repeat\_region

/note="45 copies 2 mer ac 80% conserved"

```
repeat_region 15026..15109
/note="7 copies 12 mer 82% conserved"
repeat_region 15028..15111
/note="14 copies 6 mer cacaca 82% conserved"
repeat_region 15029..15108
/note="20 copies 4 mer acac 83% conserved"
repeat_region 15273..15399
/note="L2 repeat: matches 2159..2285 of consensus"
repeat_region 15420..15721
/note="AluX repeat: matches 1..303 of consensus"
repeat_region 16333..16396
/note="2 copies 32 mer 98% conserved"
repeat_region 16575..16690
/note="L2 repeat: matches 2596..2711 of consensus"
repeat_region 18049..18169
/note="MIR repeat: matches 86..211 of consensus"
repeat_region 18312..18438
/note="MIR repeat: matches 138..250 of consensus"
repeat_region 18585..18776
/note="MER20 repeat: matches 7..217 of consensus"
repeat_region 19211..19290
/note="MIR repeat: matches 72..154 of consensus"
repeat_region 20413..20463
/note="MIR repeat: matches 90..143 of consensus"
misc_feature complement(21005..21672)
/note="match: GSS: Em:A0748384"
misc_feature complement(21187..21672)
/note="match: GSS: Em:A0776209"
misc_feature complement(21278..21663)
/note="match: GSS: Em:AQ136459"
misc_feature complement(21392..21672)
/note="match: GSS: Em:A0892114"
repeat_region 21667..21775
/note="MIR repeat: matches 73..192 of consensus"
repeat_region 23577..23646
/note="2 copies 35 mer 100% conserved"
repeat_region 24257..24591
/note="L2 repeat: matches 2336..2710 of consensus"
repeat_region 25580..25611
/note="16 copies 2 mer tc 87% conserved"
repeat_region 26333..26643
/note="Alu repeat: matches 1..310 of consensus"
repeat_region 27603..27684
/note="L2 repeat: matches 2661..2739 of consensus"
repeat_region 27823..28041
/note="MIR repeat: matches 38..242 of consensus"
repeat_region 28129..28259
/note="FLAM_C repeat: matches 1..127 of consensus"
repeat_region 31203..31356
/note="LMB5 repeat: matches 6015..6176 of consensus"
repeat_region 32224..32522
/note="AluX repeat: matches 1..298 of consensus"
repeat_region 33285..33481

Query Match 63.3%; Score 545; DB 9; Length 130435;
Best Local Similarity 100.0%; Pred. No. 4.2e-275;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 557 CCCCAACAGAACATCTTCAGACGTGACGTCATGATAGTCCAGGCTGGGCGCCCT 616
|
|
|
Db 128203 CCCCAACAGAACATCTTCAGACGTGACGTCATGATAGTCCAGGCTGGGCGCCCT 128144
|
|
|
QY 617 GCGCCCGGAGAGATCTGGGATCAGCGCCAGCTGTACGGCAGCGGGCGCATGG 676
|
|
|
Db 128143 GCGCCCGGAGAGATCTGGGATCAGCGCCAGCTGTACGGCAGCGGGCGCATGG 128084
|
|
|
QY 677 AGGGGCGCGCGCCACCTACAGGAGTTCATGGCCACTACCGGGGTCTCTTCAGC 736
|
|
|
Db 128083 AGGGGCGCGCGCCACCTACAGGAGTTCATGGCCACTACCGGGGTCTCTTCAGC 128024
|
|
|
QY 737 ACCAGCAGACGATGGGCGCGCCCTTCGTGGAGGGAGCCGGCTCCACACACACA 796
|
|
|
Db 128023 ACCAGCAGACGATGGGCGCGCCCTTCGTGGAGGGAGCCGGCTCCACACACACA 127964
|
|
|
QY 797 TCGCGCCCTAGAGAGCGGAGCCATCTGGAGCAAGAGATTAACGAAAGAGACAC 856
|
|
|
Db 127963 TCGCGCCCTAGAGAGCGGAGCCATCTGGAGCAAGAGATTAACGAAAGAGACAC 127904
|
|
|
QY 857 CTCTC 861
|
|
|
Db 127903 CTCTC 127899
|
|
|

RESULT 9
AX071267 408 bp DNA linear PAT 25-JAN-2001
LOCUS Sequence 1739 from Patent WO0102568.
DEFINITION AX071267
ACCESSION AX071267
VERSION AX071267.1 GI:12581618
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 408)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randoz,F., Kennedy,G.C., Pet,D.,
Lanson,G., Dimauc,R., Girenjakov,R., Dimauc,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
Human genes and gene expression products
Patent: WO 0102568-A 1739 11-JAN-2001;
JOURNAL CHIRON CORPORATION (US); HYSQ, INC. (US)
FEATURES
Source location/qualifiers
1..408
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 81 a 159 c 115 g 53 t
ORIGIN

Query Match 46.6%; Score 401; DB 6; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.3e-199;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Db 248 TCCCTACTGACGACGAGATGACCTGACCCACCATCTCGGTGTCAGCGGGAGGA 307
QY 471 GCGCCCACTACGACGAGGCGCCCTGACCTTCAGCTGGAGACCCGACGAGCTGGA 530
Db 308 GCGCCCACTACGACGAGGCGCCCTGACCTTCAGCTGGAGACCCGACGAGCTGGA 367
QY 531 ACTGACCGGGAGTGGTGGCGGACCCCAACAGAACCA 571
Db 368 ACTGACCGGGAGTGGTGGCGGACCCCAACAGAACCA 408

RESULT 10
AX392430/c 693 bp DNA linear PAT 23-MAR-2002
LOCUS Sequence 14 from Patent WO0216416.
ACCESSION AX392430 GI:19700746
VERSION AX392430.1
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Lee,R.T., Landschultz,K.T., Kennedy,S.P., Thompson,J.F. and
AUTHORS Turi,T.G.
TITLE diagnosis and treatment of cardiovascular conditions
JOURNAL Patent: WO 0216416-A 14 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES
source 1..693
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 101 a 205 c 237 g 149 t 1 others
ORIGIN

Query Match 35.1%; Score 302; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 2.8e-147;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 CCCCACCTTACGAGGCGCCCTGACCTTCAGCTTCGGAGACCCGAGACGAGCTGGAC 532
Db 587 CCCCACCTTACGAGGCGCCCTGACCTTCAGCTTCGGAGACCCGAGACGAGCTGGAC 528
QY 533 TGAACCGGGAGTGGTGGCGGACCCCAACAGAACCATCTTGACAGTGAAGCTGATGG 592
Db 527 TGAACCGGGAGTGGTGGCGGACCCCAACAGAACCATCTTGACAGTGAAGCTGATGG 468
QY 593 ATAGTCCAGGCTGGGCGGCGCCCTGCGCCACAGTAACTGGGATCAGCGCCACAGT 652
Db 467 ATAGTCCAGGCTGGGCGGCGCCCTGCGCCACAGTAACTGGGATCAGCGCCACAGT 408
QY 653 GCTACGCGACGCGGCGGCGGATGAGGAGGCGCGCCACCTACAGCGAGGTGATCGGCC 712
Db 407 GCTACGCGACGCGGCGGCGGATGAGGAGGCGCGCCACCTACAGCGAGGTGATCGGCC 348
QY 713 ACTACCGGGGGTCTCTCTTCAGACGACGAGACGAGGCGCCCTCTCTTGTGTGAGG 772
Db 347 ACTACCGGGGGTCTCTCTTCAGACGACGAGACGAGGCGCCCTCTTGTGTGAGG 288
QY 773 GG 774
Db 287 GG 286

RESULT 11
AX011709 812 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 107 from Patent WO9955858.
ACCESSION AX011709
VERSION AX011709.1 GI:9998233
KEYWORDS human.
SOURCE

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 812)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pillarsky,C.
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue
JOURNAL Patent: WO 9955858-A 107 04-NOV-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES
source 1..812
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 157 a 272 c 237 g 146 t
ORIGIN

Query Match 31.5%; Score 271; DB 6; Length 812;
Best Local Similarity 99.7%; Pred. No. 5.6e-131;
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 173 TCACGTGCTCTGACGACCTACCAAGCTGTCTGACGCTCTTCATCAGCCGACAGCC 232
Db 426 TCACGTGCTCTGACGACCTACCAAGCTGTCTGACGCTCTTCATCAGCCGACAGCC 485
QY 233 AGGGCGGAGAGAGAGATGCTCTGCTCAGAAAGATGCTGTGGCCCTCGGAGAGCA 292
Db 486 AGGGCGGAGAGAGAGATGCTCTGCTCAGAAAGATGCTGTGGCCCTCGGAGAGCA 545
QY 293 CAGTGTGAGGACGAGGATTCACAGCGGAGGAGTACGCGCCGCTGGGCGACCGAGC 352
Db 546 CAGTGTGAGGACGAGGATTCACAGCGGAGGAGTACGCGCCGCTGGGCGACCGAGC 605
QY 353 GCTTGGCGGCGCGCCCTTTCGCGGAGGAGCGCTTCACCGCTTCAGGCCACCTATTC 412
Db 606 GCTTGGCGGCGCGCCCTTTCGCGGAGGAGCGCTTCACCGCTTCAGGCCACCTATTC 665
QY 413 CGTACCTGACGACGAGATGACCTGCACCCACCATCTGCTGTCAGAGGGAGAGGC 472
Db 666 CGTACCTGACGACGAGATGACCTGCACCCACCATCTGCTGTCAGAGGGAGAGGC 725
QY 473 CCCCACCTTACGAGGCGCCCTG 494
Db 726 CCCCACCTTACGAGGCGCCCTG 747

RESULT 12
HSJ10591L/ 150224 bp DNA linear PRI 24-FEB-2001
LOCUS Human DNA sequence from clone R95-10591L on chromosome
20q13.2-13.33 contains the 5' end of the TMEPA1 (PMEPA1) gene
encoding an androgen induced 1b transmembrane protein, ESTs, STSs,
GSSs and two CpG islands, complete sequence.
ACCESSION AL121913
VERSION AL121913.4 GI:7161781
KEYWORDS HMG; CpG island; PMEPA1; TMEPA1; transmembrane protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 150224)
AUTHORS Skuce,C.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Mar 6, 2000 this sequence version replaced gi:7007305.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with

```



repeat_region	/note="MIR repeat: matches 107. .243 of consensus"
	19070. .19291
repeat_region	/note="MIR repeat: matches 35. .261 of consensus"
	19954. .20005
repeat_region	/note="L2 repeat: matches 2358. .2411 of consensus"
	20264. .20432
repeat_region	/note="L2 repeat: matches 2581. .2750 of consensus"
	20530. .21282
repeat_region	/note="L1MD1 repeat: matches 5595. .6341 of consensus"
	21980. .22192
repeat_region	/note="L2 repeat: matches 2322. .2562 of consensus"
	21102. .23203
repeat_region	/note="MIR repeat: matches 117. .218 of consensus"
	23231. .23438
repeat_region	/note="MER20 repeat: matches 3. .218 of consensus"
	23439. .23512
repeat_region	/note="MIR repeat: matches 182. .257 of consensus"
	24460. .24580
repeat_region	/note="L1MD1 repeat: matches 6102. .6221 of consensus"
	24752. .25042
repeat_region	/note="L2 repeat: matches 2386. .2709 of consensus"
	25208. .25295
repeat_region	/note="MIR repeat: matches 61. .155 of consensus"
	26450. .26485
repeat_region	/note="18 copies 2 mer by 97% conserved"
	29374. .29501
repeat_region	/note="MIR repeat: matches 55. .190 of consensus"
	29550. .29728
repeat_region	/note="MIR repeat: matches 69. .248 of consensus"
	30974. .31029

Query Match	18.1%;	Score 156;	DB 9;	Length 150224;
Best Local Similarity	100.0%;	Prod. No. 1.8e-70;		
Matches 156;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 110	CGGAGCTGAGATTGTTGTCAGATCATCATCATCGTGGTGGTGTGATGTGATGTGTGGTGG	169		
Db 2599	CGGAGCTGAGATTGTTGTCAGATCATCATCATCGTGGTGGTGTGATGTGATGTGTGGTGG	2540		
QY 170	TGATCAGCTGGCTGCTGTGAGCCACTCAAGCTGTGTGACAGGTCCTTCATCAGCCGGGCACA	229		
Db 2539	TGATCAGCTGGCTGCTGTGAGCCACTCAAGCTGTGTGTGACAGGTCCTTCATCAGCCGGGCACA	2480		
QY 230	GCCAGGGGGCGAGAGAGAGAGAGATGCCCTGTCTCTCAG	265		
Db 2479	GCCAGGGGGCGAGAGAGAGAGATGCCCTGTCTCTCAG	2444		
RESULT 13				
LOCUS	AX199565	51 bp	DNA	linear
DEFINITION	Sequence 495 from Patent WO0151670.			
ACCESSION	AX199565			
VERSION	AX199565.1	GI:15389996		
KEYWORDS	human.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 51)			
TITLE	Shinketsu, R.A. and Leach, M.D.			
JOURNAL	Nucleic acids containing single nucleotide polymorphisms and methods of use thereof			
	Patent: WO 0151670-A 495 19-JUL-2001;			
FEATURES	Curagen Corporation (US)			
source	Location/Qualifiers			
	1..51			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
misc_feature	26			
	/note="1 of 2 allelic variants (496 is other entry)			
BASE COUNT	12 a	16 c	18 g	5 t

ORIGIN	Query Match	5.9%	Score 51	DB 6	Length 51
	Best Local Similarity	100.0%	Pred. No. 3e-15		
	Matches 51	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	206 CACGGTCCTTCATCAGCCGGCACAGCCAGCGGGCGGAGAGAGACATGCCCC			256	
Db	1 CACGGTCCTTCATCAGCCGGCACAGCCAGCGGGCGGAGAGAGACATGCCCC			51	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	AUTHORS	TITLE	JOURNAL	FEATURES	CDs	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches	OY	Db	RESULT 15	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428					
AF220208	Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.	AF220208				Mus musculus.	1	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		Biochem. J.	351 Pt 3, 557-565 (2000)	20498735						146 a	234 c	176 g	95 t				AX392428	Sequence 12 from Patent WO0216416.	AX392428			</		



REFERENCE 1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and  
Tui, T.G.  
TITLE Diagnosis and treatment of cardiovascular conditions  
JOURNAL Patent: WO 0216416-A 12 28-FEB-2002;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)  
FEATURES  
source 1. 878  
location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
20. .844  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD28011.1"  
/db\_xref="GI:19700745"  
/translation="MGVNGTAATAAAGQPNVSCACNCORSLEPPSMETLEFVOIVIV  
VVMVMVMTICLSHYKLSARSPISRHSQARRDDGLSEGLPSSSTVSGMPEP  
QVAPRPRTDLAVPEFIQSRFOPTVYLOHEIALPTISLSDGEEPPYOGPCTLO  
LRDPEOLELNRESYRAPNRTIFDSDLIDSTMIGGPGCPSSNSGISATCYSSGGME  
GPPPTYSVIGHYSSSFQHOOSNGRSSLLECTRLHSHIAPLENKEKEKOKGHP.L"  
BASE COUNT 179 a 305 c 247 g 147 t  
ORIGIN  
Query Match 5.1%; Score 44; DB 6; Length 878;  
Best Local Similarity 100.0%; Pred. No. 1.5e-11;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 463 GGGGAGAGCCCCCACCCTACCGAGGCCCCCTGCACCTCCAGCT 506  
|||||  
Db 458 GGGGAGAGCCCCCACCCTACCGAGGCCCCCTGCACCTCCAGCT 501  
|||||

Search completed: March 18, 2003, 07:55:51  
Job time : 2101 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 07:12:45 ; Search time 210 Seconds  
(without alignments)  
9233.193 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861  
Sequence: 1 atgcacgcgtgatgagggc.....aacagaagacacccctc 861

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N.Geneseq-101002.\*  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	861	100.0	1321	24	ABK12137 Human CDNA encodin
2	752	87.3	1140	24	ABK92120 Prostate cancer-as
3	740	85.9	1066	22	AA157868 Human polynucleoti
4	701	81.4	969	21	AAA75151 CDNA encoding a hu
5	701	81.4	1061	21	AAA47429 Sequence encoding
6	696	80.8	969	21	AAA75163 CDNA clone encodin
7	650	75.5	969	21	AAA75164 CDNA clone encodin
8	650	75.5	969	21	AAA75165 CDNA clone encodin
9	401	46.6	408	22	AAF65983 Novel human polynu

C	10	346	40.2	1069	22	AA159654 Human polynucleoti
C	11	302	35.1	693	20	ABK12143 Human MIVR-1 homol
C	12	271	31.5	812	24	AA252964 Human prostate tum
C	13	229	26.6	254	21	AAA41265 Human secreted exp
C	14	106	12.3	1879	23	AA584503 DNA encoding novel
C	15	51	5.9	51	22	AAH89714 Human coding sequ
C	16	49	5.7	60	24	ABN40872 Human spliced tran
C	17	44	5.1	878	24	ABK12142 Mouse CDNA encodin
C	18	44	5.1	1713	21	AAA75152 CDNA encoding a mu
C	19	44	5.1	1713	21	AAA75166 CDNA clone encodin
C	20	44	5.1	1713	21	AAA75167 CDNA clone encodin
C	21	44	5.1	1713	21	AAA75168 CDNA clone encodin
C	22	44	4.8	475	24	ABK12144 Human MIVR-1 homol
C	23	32	3.7	577	22	ABA50270 Human breast cell
C	24	32	3.7	577	22	ABA68210 Human foetal liver
C	25	32	3.7	577	22	ABA35222 Probe #13688 for g
C	26	32	3.7	577	22	AAK16592 Human brain expres
C	27	32	3.7	577	22	AAK42346 Human bone marrow
C	28	32	3.7	577	22	AAI23118 Probe #13051 for g
C	29	32	3.7	577	22	AAI48425 Probe #17111 used
C	30	32	3.7	577	22	AAI08772 Probe #8763 used t
C	31	32	3.7	577	24	ABSI6402 Human genome-deriv
C	32	32	3.7	1964	22	ABA45134 Human breast cell
C	33	32	3.7	1964	22	ABA55612 Human foetal liver
C	34	32	3.7	1964	22	ABA25310 Human foetal liver
C	35	32	3.7	1964	22	AAK03841 Human brain expres
C	36	32	3.7	1964	22	AAK29307 Human bone marrow
C	37	32	3.7	1964	22	AAI13901 Probe #3834 for ge
C	38	32	3.7	1964	22	AAI35267 Probe #3953 used t
C	39	32	3.7	1964	22	AAI03769 Probe #3760 used t
C	40	32	3.7	1964	24	ABSO3866 Human genome-deriv
C	41	31	3.6	10437	22	AA541709 Genomic sequence #
C	42	31	3.6	12620	22	ABAI6691 Human nervous syst
C	43	30	3.5	2109	23	AA581477 DNA encoding novel
C	44	29	3.4	401	22	AA545299 Human breast cell
C	45	29	3.4	401	22	ABA55787 Human foetal liver

## ALIGNMENTS

RESULT 1	ABK12137	standard; cDNA; 1321 bp.
ID	ABK12137	
AC	ABK12137	
XX		
DT	05-JUN-2002	(first entry)
XX		
DE	Human CDNA encoding Mechanically Induced Vascular Receptor 1, MIVR-1.	
XX		
KW	Human; ss; gene; MIVR-1; Mechanically Induced Vascular Receptor 1;	
KW	cytostatic; cardi-ant; cerebroprotective; antiarteriosclerotic;	
KW	cardiac cell; anti-apoptotic; vascular endothelial cell;	
KW	cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;	
KW	heart failure.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	413..1276
FT		/*tag= a
FT		/product= "MIVR-1"
FT		/note= "This region is specifically claimed in claim 3"
XX		
XX	MO200216416-A2.	
XX		
PN	28-FEB-2002.	
XX		
PD		
XX	21-AUG-2001; 2001WO-US26089.	
XX		
XX	22-AUG-2000; 2000US-227159P.	
XX		



CC compris contacting a biological sample from the patient with  
CC prostate cancer-associated polynucleotides (designated PC genes) that  
CC selectively hybridise to a sequence that is at least 80% identical  
CC to them. The prostate cancer-associated polynucleotide sequences  
CC are differentially expressed in prostate tumour tissue or in  
CC prostate cancer and are derived from the tissues of various  
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs)  
CC The methods of the invention are useful for diagnosing and treating  
CC prostate cancer in mammals. The prostate cancer-associated genes are  
CC useful for diagnosing or treating prostate cancer, as well as for  
CC identifying modulators of prostate cancer or agents that inhibit  
CC prostate cancer. The nucleic acid sequences are particularly useful  
CC in gene therapy, as a vaccine or in antisense applications.  
CC ABR62115-ABR62263 represent prostate cancer-associated polynucleotide  
CC sequences.  
50 Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 other;  
5X

Query Match	87.38;	Score 752;	DB 24;	Length 1140;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 752; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

OY	110	CGGACCTGGAGTTTGTTCATCATCATCAATCAGTGSTGSGTGAATGATGATGATGATG	169
Dp	100	CGGACTGTGAGTTTCTTCATGATCATCATCATGTTGGTGTGATGATGATGATGATG	159
OY	170	TGATACGCTGCCCTGCTGAGACCCTACAAGCTGTCTGCAGCGTCTTCATCAGCCGGCAC	229
Dp	160	TGATACGCTGCCCTGCTGAGCCACTCAAACTGTCTGCAGCGTCTTCATCAGCCGGCAC	219
OY	230	GCCAGGGCCGAGAGAGAAGATGCCCCCTGTCTCAGAAAGATGCCGTGTGGCCCTCGAGA	289
Dp	220	GCCAGGGCCGAGAGAGAAGATGCCCTGTCTCTCAGAAAGATGCCGTGTGGCCCTCGAGA	279
OY	290	GCACAGTGTCAAGGCACAACGAAATCCCAGAGCCCGCAGGTCTACGCCCCCGCTGGCCACCG	349
Dp	280	GCACAGTGTCAAGGCACAACGAAATCCCAGAGCCCGCAGGTCTACGCCCCCGCTGGCCACCG	339
OY	350	ACCGCTGTGGCGTGGCGGCCCTTGCCCTCAGACGGAGAGCGTTCACCGCGTTCACGCCACCT	409
Dp	340	ACCGCTGTGGCGTGGCGGCCCTTGCCCTCAGACGGAGAGCGTTCACCGCGTTCACGCCACCT	399
OY	410	ATCCGTACCTGCAGCAGCAGATGCATGCATGCCACCCACATCTCGCTGTGCAGACGGGAGG	469
Dp	400	ATCCGTACCTGCAGCAGCAGATGCATGCATGCCACCCACATCTCGGTGTGCAGACGGGAGG	459
OY	470	AGCCCCCAACCTACACAGGGCCCCCTGCACCCCTTCAGACTTGGGAGCCCGACAGCACTGG	529
Dp	460	AGCCCCCAACCTACACAGGGCCCCCTGCACCCCTTCAGACTTGGGAGCCCGACAGCACTGG	519
OY	530	AACITGAACCGGAGTGCGGTGCAGCACCACCCCAAACAGAACATCTTCGACAGTGAACCTGA	589
Dp	520	AACITGAACCGGAGTGCGGTGCAGCACCACCCCAAACAGAACATCTTCGACAGTGAACCTGA	579
OY	590	TGGATAGTGCAGAGCTGGGGGCCCTGCCTGCCTCCAGCAGTAATCTGGGCAATCAGGCCCA	649
Dp	580	TGGATAGTGCAGAGCTGGGGGCCCTGCCTGCCTCCAGCAGTAATCTGGGCAATCAGGCCCA	639
OY	650	CGTGTACGGCAGCGCGGGGGGCGCATATGAGAGGGGCCCGCCACCTTAACAGCAGATCATG	709
Dp	640	CGTGTACGGCAGCGCGGGGGGCGCATATGAGAGGGGGCCCGCCACCTTAACAGCAGATCATG	699
OY	710	GCCACTACCCGGGGTCTCTCTTCAGAGCAACACAGCAGAGCAGTGGGGCGCCCTCTTGGCTGG	769
Dp	700	GCCACTACCCGGGGTCTCTCTTCAGAGCAACAGCAGAGCAGTGGGGCGCCCTCTTGGCTGG	759
OY	770	AGGGGACCCGGCTCCACACACACACATTCGCGCCCTTAGAGAGCGCAGGCATCTGGAGCA	829
Dp	760	AGGGGACCCGGCTCCACACACACACATTCGCGCCCTTAGAGAGCGGAGCATCTGGAGCA	819
OY	830	AAGAGAAGGATTAAACAGAAAAGCAACCTCTC	861
Dp	820	AAGAGAAGGATTAAACAGAAAAGCAACCTCTC	851

```

RESULT 3
AA157868
ID   AA157868 standard; cDNA; 1066 BP

```

DT 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 71.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation  
KW leukaemia; ss.

OS Homo sapiens.

PN W0200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 09-JUL-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0653450.

PR 19-OCT-2000; 2000US-0693036.

XX XX

XX  
X

PI wang J, wang Z, wehrman T, xu C, xue AJ, yang Y, zhang J;

XX  
XX

EEDT  
0001  
0002  
0003  
0004  
0005  
0006  
0007

DR P-PSDB; AAM38712.

Novel nucleic acids and polypeptides, useful for treating disorders

[illegible]

XX  
XX  
XX

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC system, such as peripheral nervous injuries, peripheral neuropathy and localized neuroathic and central nervous system diseases such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis and Charcot-Marie-Tooth. Other uses include the

CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin chemotactic/chemokinetic activity, chemotactic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, antibodies and inflammation, leukemias and

CC C.N.S disorders.  
CC Note: The sequence data for this patient did not form part of the stratified

CC specification.  
XX

sequence 1066 BP; 225 A; 348 C; 325 G; 168 T; 0 other;

Query Match	85.9%	Score 740	DB 22	Length 1066
Best Local Similarity	100.0%	Pred. No.	0	
Matches 740	Conservative 0	Mismatches 0	Indels 0	Gaps 0



Db 130 GCCAGGGGGGAGAGAGATGCCCTGTCCTCAGAGAGATGCTGTGGCCCTCGGAGA 189  
QY 290 GCACAGTGTAGCGACGGAATCCAGAGCCGAGGTCTACGGCCCGCTCGGCCACCG 349  
Db 190 GCACAGTGTAGCGACGGAATCCAGAGCCGAGGTCTACGGCCCGCTCGGCCACCG 249  
QY 350 ACCGCGCGGGCGTGGCCCTTGGCCCGAGGGGAGCGCTTCGACCGCTTCAGACCCACT 409  
Db 250 ACCGCGCGGGCGTGGCCCTTGGCCCGAGGGGAGCGCTTCGACCGCTTCAGACCCACT 309  
QY 410 ATCCGTAAGTGCAGCAGCAGAGATGACCTGGCCACCACCATCTGCTGTGAGCGGGAGG 469  
Db 310 ATCCGTAAGTGCAGCAGCAGAGATGACCTGGCCACCACCATCTGCTGTGAGCGGGAGG 369  
QY 470 AGCCCGCCACCTTACAGAGGCGCCCTGCACCTCCAGCTTGGGGAGCCCGAGCAGAGCTGG 529  
Db 370 AGCCCGCCACCTTACAGAGGCGCCCTGCACCTCCAGCTTGGGGAGCCCGAGCAGAGCTGG 429  
QY 530 AACTGAACCGGGAGTGGTGGCGGCAACCCCAACAGAACATCTTGCAGAGTGAAGCTGA 589  
Db 430 AACTGAACCGGGAGTGGTGGCGGCAACCCCAACAGAACATCTTGCAGAGTGAAGCTGA 489  
QY 590 TGGATAGTGCAGCAGCTGGGGCGCCCTGCCCCCAGCAGTAACTCGGCGATCAGCGCA 649  
Db 490 TGGATAGTGCAGCAGCTGGGGCGCCCTGCCCCCAGCAGTAACTCGGCGATCAGCGCA 549  
QY 650 CGTGTACGCGAGCGCGCGGCGCATGAGGGGCGCCGCCACCTACAGCAGAGTCAATCG 709  
Db 550 CGTGTACGCGAGCGCGCGGCGCATGAGGGGCGCGCCGCCACCTACAGCAGAGTCAATCG 609  
QY 710 GCCACTACCGGGGGTCTCTCTTCAGAGCAGCAGAGCAGTGGGGCGCCCTCTCTGTGG 769  
Db 610 GCCACTACCGGGGGTCTCTCTTCAGAGCAGCAGAGCAGTGGGGCGCCCTCTCTGTGG 669  
QY 770 AGGGGAGCCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGCATCTGAGCA 829  
Db 670 AGGGGAGCCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGCATCTGAGCA 729  
QY 830 AAGAGAAGATTAAGAAAGAGACCCCTCTC 861  
Db 730 AAGAGAAGATTAAGAAAGAGACCCCTCTC 761

RESULT 5  
AAAA7429  
ID AAAA7429 standard; DNA; 1061 BP.  
XX  
AC AAAA7429;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Sequence encoding human neuron-associated protein.  
XX  
KW Neuron associated protein; NEUAP; neurological disorder; epilepsy;  
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
KW Alzheimer's disease; Pick's disease; Huntington's disease;  
KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;  
KW cerebral palsy; muscular dystrophy; central nervous system; CNS;  
KW peripheral nervous system; PNS; myopathy; schizophrenia;  
KW acanthosis; keratosis; arteriosclerosis; atherosclerosis; burstitis;  
KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
KW myelofibrosis; paroxysmal nocturnal hemoglobinuria; cancer;  
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KW AIDS; Addison's disease; adult respiratory distress syndrome;  
KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KW Werner syndrome; trauma; human; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 101..859  
FT /\*tag- a

FT /product= Neuron associated protein  
XX  
PN W0200034477-A2.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99NM-US30408.  
XX  
PR 11-DEC-1998; 98US-0210083.  
PR 11-DEC-1998; 98US-9123456.  
PR 09-FEB-1999; 99US-0119365.  
PR 16-MAR-1999; 99US-0124687.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-Young J, Yang J;  
PI Lu DAM, Azimzal Y.  
XX  
DR WPI: 2000-423423/36.  
DR P-PSDB; AAB01388.  
XX  
PT New human neuron-associated proteins and polynucleotides encoding them,  
PT useful for diagnosis, treatment and prevention of cell proliferative  
PT disorders including cancer, neuronal and neurological disorders  
XX  
PS Claim 9; Page 136; 145pp; English.  
XX  
CC Human neuron-associated proteins (NEUAP) can be used for  
CC treating or preventing a disorder associated with decreased  
CC expression or activity of NEUAP. Antagonists of NEUAP are useful for  
CC treating or preventing disorder associated with increased expression  
CC or activity of NEUAP. NEUAP or their fragments or derivatives are  
CC useful for treating neurological disorder such as epilepsy, ischemic  
CC cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's  
CC disease, Pick's disease, Huntington's disease, dementia and  
CC Parkinson's disease. NEUAPs are also useful for treating other  
CC demyelinating diseases, bacterial and viral meningitis, prion  
CC diseases including kuru, Creutzfeldt-Jakob disease, nutritional and  
CC metabolic diseases of the nervous system, neurofibromatosis, other  
CC developmental disorders of the central nervous system, cerebral  
CC palsy, neuroskeletal disorders, autonomic nervous system disorders,  
CC cranial nerve disorders, spinal cord diseases, muscular dystrophy and  
CC other neuromuscular disorders, peripheral nervous system disorders,  
CC inherited, metabolic, endocrine, and toxic myopathies, mental  
CC disorders including mood, anxiety and schizophrenic disorders, a cell  
CC proliferative disorder such as actinic keratosis, arteriosclerosis,  
CC atherosclerosis, burstitis, cirrhosis, hepatitis, mixed connective  
CC tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal  
CC haemoglobinuria, cancers of the adrenal gland, bladder, bone,  
CC bone marrow, brain, breast, cervix, and an autoimmune/inflammatory  
CC disorder such as acquired immunodeficiency syndrome (AIDS), Addison's  
CC disease, adult respiratory distress syndrome, allergies, ankylosing  
CC spondylitis, amyloidosis, anemia, asthma, Werner syndrome,  
CC complications of cancer, hemodialysis, and extracorporeal circulation,  
CC viral, bacterial, fungal parasitic, protozoal, and helminthic  
CC infections, and trauma. This sequence was given the Incyte ID no.  
CC 1671288CB1.  
XX  
SQ Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 other;  
XX

Query Match 81.4%; Score 701; DB 21; Length 1061;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 751; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 110 CGGAGCTGGAAGTTTGTTCAGATCATCATCGTGGTGATGATGCTGTGGTGG 169  
Db 105 CGGAGCTGGAAGTTTGTTCAGATCATCATCGTGGTGATGATGCTGTGGTGG 164  
QY 170 TGATCAGCTGCTGTGAGCAGCAGTCAAGCTGTGCGAGGCTCTTATCAGCGGCA 229  
Db 165 TGATCAGCTGCTGTGAGCAGCAGTCAAGCTGTGCGAGGCTCTTATCAGCGGCA 224  
QY 230 GCCAGGGGGGAGAGAGATGCCCTGTCCTCAGAAAGATGCTGTGGCCCTCGGAGA 289

```

Db 225 GCGAGGGGCGAGAGAGAGATGCTGCTCAGAGAGATGCTGCGCCCTGGAGAGA 284
QY 290 GCACAGTGTAGAGCAGAGATCCAGAGCGAGGTCAGCGCCCGCTCGCCACCG 349
Db 285 GCACAGTGTAGAGCAGAGATCCAGAGCGAGGTCAGCGCCCGCTCGCCACCG 344
QY 350 ACCGCTGCGCGCTGCGCCCTGCGCCAGCGGAGCGCTTCACACCGTTTCAGCCACT 409
Db 345 ACCGCTGCGCGCTGCGCCCTGCGCCAGCGGAGCGCTTCACACCGTTTCAGCCACT 404
QY 410 ATCCGTAACCTGACAGCAGATGACCTGACACCCACATCTGCTGTACAGAGGGAGG 469
Db 405 ATCCGTAACCTGACAGCAGATGACCTGACACCCACATCTGCTGTACAGAGGGAGG 464
QY 470 AGCCGCCACCTTACAGAGGCGCTGACACCTGACAGTTGGGAGCCCGACACAGCTG 529
Db 465 AGCCGCCACCTTACAGAGGCGCTGACACCTGACAGTTGGGAGCCCGACACAGCTG 524
QY 530 AACTGAACCGGAGAGTGGTGGCGGCAACCCCAAGAACATCTTGCAGAGTACCTGA 589
Db 525 AACTGAACCGGAGAGTGGTGGCGGCAACCCCAAGAACATCTTGCAGAGTACCTGA 584
QY 590 TGGATAGTGTGACAGCGGCGCGCCCTGCGCCCGACAGTAACTGGGCGATCAGCGCA 649
Db 585 TGGATAGTGTGACAGCGGCGCGCCCTGCGCCCGACAGTAACTGGGCGATCAGCGCA 644
QY 650 CGTGCTACGCGACGCGCGGCGCGATGAGAGGCGCGCCCGACCTACAGCGAGGTATG 709
Db 645 CGTGCTACGCGACGCGCGGCGCGATGAGAGGCGCGCCCGACCTACAGCGAGGTATG 704
QY 710 GGCACATACCCGGGCTCTCTCTTCCAGACACAGCAGAGCTGAGGCGCGCTCTCTG 769
Db 705 GGCACATACCCGGGCTCTCTCTTCCAGACACAGCAGAGCTGAGGCGCGCTCTCTG 764
QY 770 AGGGGACCGCGGCTCCACACACACATGCGCGCTTACAGAGCGAGCATCTGAGAGA 829
Db 765 AGGGGACCGCGGCTCCACACACACATGCGCGCTTACAGAGCGAGCATCTGAGAGA 824
QY 830 AAGAGAGAGATTAACAGAAAGACACCCCTCTC 861
Db 825 AAGAGAGAGATTAACAGAAAGACACCCCTCTC 856

RESULT 6
AAA75163
ID AAA75163 standard; cDNA; 969 BP.
XX
XX AAA75163;
XX
XX 15-JAN-2001 (first entry)
XX
XX cDNA clone encoding a human TANGO 261 polypeptide.
DE
XX
XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;
XX cellular proliferation; cellular differentiation; cellular adhesion;
XX von Willebrand factor-associated disorder; cell trafficking; cancer;
XX hematopoietic associated disease; atelectasis; pulmonary congestion;
XX oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
XX intestinal disorder; spleen associated disease; renal disorder;
XX cardiovascular disorder; ischemic heart disease; hydrocephalus;
XX brain herniation; iatrogenic disease; inflammation; meningitis;
XX Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;
XX multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 6..764
XX FT /*tag= a
XX FT /product= "TANGO 261"
XX
XX W0200052022-A1.

```

```

XX
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-US05226.
XX
PR 01-MAR-1999; 99US-0122458.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CJ;
XX
DR WPI: 2000-579269/54.
XX
P-PSDB: AAB18461.
XX
PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -
XX
PS Disclosure; Page -: 175pp; English.
XX
CC AAA75163-65 encode human TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammation,
CC bacterial and viral meningitis, Alzheimer's disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.
XX
XX
SO Sequence 969 BP; 211 A; 317 C; 293 G; 148 T; 0 other;
XX
XX
XX Query Match 80.8%; Score 696; DB 21; Length 969;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 746; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 115 CTGAGTTTGTTCAGATCATCATCATGCTGCTGCTGATGATGATGATGATGATC 174
Db 15 CTGAGATTGTTCAGATCATCATCATGCTGCTGCTGATGATGATGATGATGATC 74
QY 175 ACGTCCCTGCTGAGCCTACACACTGCTGACAGGCTCTTCATCAGCCGACAGCAG 234
Db 75 ACGTCCCTGCTGAGCCTACACACTGCTGACAGGCTCTTCATCAGCCGACAGCAG 134
QY 235 GGGGAGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
Db 135 GGGGAGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194
QY 295 GTGTGAGCAAGGATCCAGAGCCGACAGTCTAGCCCGCTGCGCCACAGCCG 354
Db 195 GTGTGAGCAAGGATCCAGAGCCGACAGTCTAGCCCGCTGCGCCACAGCCG 254
QY 355 CTGGCGCTGCGCGCTTCCGCGCGAGCGGAGCGCTTCCACCGCTTCCAGCCACTATCG 414
Db 255 CTGGCGCTGCGCGCTTCCGCGCGAGCGGAGCGCTTCCACCGCTTCCAGCCACTATCG 314
QY 415 TACCTGAGCAGAGATGACCTGCGCACCCACATCGCTGCTGACAGCGGAGGAGCCG 474

```





Oy 650 CTTGCTACGGGCGGCGGCGCATGAGAGGGGCGGCGCCACCTACAGCGAGTCTATCG 709  
 |||||||  
 Db 550 CGTGTACGGGCGGCGGCGCGCATGAGAGGGGCGGCGCCACCTACAGCGAGTCTATCG 609  
 |||||||  
 Oy 710 GCCACTACCGGGGCTCTCTTCACAGCACAGAGAGAGTGGGCGGCTCTCTTGTGG 769  
 |||||||  
 Db 610 GCCACTACCGGGGCTCTCTTCACAGCACAGAGAGAGTGGGCGGCTCTCTTGTGG 669  
 |||||||  
 Oy 770 AGGGAGCCGGGCTCCACACACACATCGGCGGCGCTTAGAGAGCGGACCATCTGGAGCA 829  
 |||||||  
 Db 670 AGGGAGCCGGGCTCCACACACACATCGGCGGCGCTTAGAGAGCGGACCATCTGGAGCA 729  
 |||||||  
 Oy 830 AAGAGAGGATTAACAGAAAGAGACACCTCTC 861  
 |||||||  
 Db 730 AAGAGAGGATTAACAGAAAGAGACACCTCTC 761

## RESULT 8

AAA75165  
 ID AAA75165 standard; cDNA; 969 BP.  
 XX

AC AAA75165;

DT 15-JAN-2001 (first entry)

DE cDNA clone encoding a human TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Homo sapiens.

FH key Location/Qualifiers

FT CDS 6..764  
 FT /tag= a  
 FT /product= "TANGO 261"

PN WO200052022-A1.

PD 08-SEP-2000.

PF 01-MAR-2000; 2000WO-US05226.

PR 01-MAR-1999; 99US-0122458.

PA (MILL-) MILLENNIUM PHARM INC.

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

DR WPI; 2000-579269/54.

DR P-PSDB; AAB18463.

PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -

PS Disclosure; Page -: 175pp; English.

XX AAA75163-65 encode human TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,

CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC modulate the proliferation, differentiation, and/or function of cells  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, atelectasis,  
 CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral oedema,  
 CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 CC bacterial and viral meningitis, Alzheimer's Disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC note: the present sequence does not appear in the specification, it was  
 CC created using information provided.

SX Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;

Query Match 75.5%; Score 650; DB 21; Length 969;  
 Best Local Similarity 99.7%; Pred. No. 5.8e-295;

Matches 750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 110 CGAGCTGAGTGTTCAGATCATCATCATGCTGGTGATGATGATGATGATGATG 169  
 |||||||  
 Db 10 CGAGCTGAGTGTTCAGATCATCATCATGCTGGTGATGATGATGATGATGATG 69  
 |||||||  
 Oy 170 TGAATCAGTCCCTGCTGAGACCACTACAGCTGTGACAGGCTCTTCATCAGCGGCA 229  
 |||||||  
 Db 70 TGAATCAGTCCCTGCTGAGACCACTACAGCTGTGACAGGCTCTTCATCAGCGGCA 129  
 |||||||  
 Oy 230 GCCAGGCGGAGAGAGAGATGCTCTCTCAGAAAGATGCTGTGGCTCTGAGAGA 289  
 |||||||  
 Db 130 GCCAGGCGGAGAGAGAGATGCTCTCTCAGAGAGATGCTGTGGCTCTGAGAGA 189  
 |||||||  
 Oy 290 GCACAGTGTCAAGCAAGCAATCCAGAGCGGAGCTGTACGCGGCGGCTGTGCGCCACCG 349  
 |||||||  
 Db 190 GCACAGTGTCAAGCAAGCAATCCAGAGCGGAGCTGTACGCGGCGGCTGTGCGCCACCG 249  
 |||||||  
 Oy 350 ACCGCTGTGCGGCTGTGCGGCTGTGCGGAGCGGAGCGCTTCCAGCGCCACT 409  
 |||||||  
 Db 250 ACCGCTGTGCGGCTGTGCGGCTGTGCGGAGCGGAGCGCTTCCAGCGCCACT 309  
 |||||||  
 Oy 410 ATCGCTACCTGAGCAGAGATGACCTGCAACCCACCATCTCGCTGACAGCGGAGG 469  
 |||||||  
 Db 310 ATCGCTACCTGAGCAGAGATGACCTGCAACCCACCATCTCGCTGACAGCGGAGG 369  
 |||||||  
 Oy 470 AGCCCGCACCTTACAGGAGCGGCGGCTGCAACCTGCAAGCTTGGGAGCGAGCAGCTGG 529  
 |||||||  
 Db 370 AGCCCGCACCTTACAGGAGCGGCGGCTGCAACCTGCAAGCTTGGGAGCGAGCAGCTGG 429  
 |||||||  
 Oy 530 AACTGACCGGGAGTCCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 589  
 |||||||  
 Db 430 AACTGACCGGGAGTCCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 489  
 |||||||  
 Oy 590 TGGATAGTGCAGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 649  
 |||||||  
 Db 490 TGGATAGTGCAGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 549  
 |||||||  
 Oy 650 CGTCTACGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 709  
 |||||||  
 Db 550 CGTCTACGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 609  
 |||||||  
 Oy 710 GCCACTACCGGGGCTCTCTTCACAGCACAGAGAGAGTGGGCGGCGCTCTTGTGG 769  
 |||||||  
 Db 610 GCCACTACCGGGGCTCTCTTCACAGCACAGAGAGAGTGGGCGGCGCTCTTGTGG 669  
 |||||||  
 Oy 770 AGGGAGCCGGGCTCCACACACATCGGCGGCGCTTAGAGAGCGGACCATCTGGAGCA 829  
 |||||||  
 Db 670 AGGGAGCCGGGCTCCACACACATCGGCGGCGCTTAGAGAGCGGACCATCTGGAGCA 729  
 |||||||







PT responsible for genetic disorders or other traits and to assess

Isolated human polynucleotides containing single nucleotide

PT Infection and diabetes -

PS Claim 1, Page 246; 475pp; English.

CC The present invention relates to human nucleic acids containing single  
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
 CC paternity tests, and to aid in the treatment of diseases associated with  
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
 CC glomerulonephritis, haemolytic anaemia, thrombocytopenia, arthritis,  
 CC meningitis, muscular disorders, dementia, neurological diseases, tubercu-  
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
 CC autoimmunity. The present sequence is a polymorphism-containing  
 CC oligonucleotide fragment of the invention.

XX  
 SQ Sequence 51 BP; 12 A; 16 C; 18 G; 5 T; 0 other;

Query Match 5.9%; Score 51; DB 22; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-14;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 CACGGTCTTCATCAGCCGACAGCCAGGGGCGAGAGAGAGATGCC 256  
 DB 1 CACGGTCTTCATCAGCCGACAGCCAGGGGCGAGAGAGAGATGCC 51

Search completed: March 18, 2003, 07:20:34  
 Job time : 215 secs



**THIS PAGE BLANK (USPTO)**

1	23	2.7	377	2	US-08-332-766A-1	Sequence 1, Appl1
2	21	2.4	633	1	US-08-234-783-1	Sequence 1, Appl1
3	21	2.4	633	1	US-08-456-907-1	Sequence 1, Appl1
4	21	2.4	633	5	PCT-US95-05523-1	Sequence 1, Appl1
5	21	2.4	1337	1	US-08-165-1150-3	Sequence 3, Appl1
6	21	2.4	4659	4	US-08-264-578-10	Sequence 10, Appl1
7	21	2.4	8147	4	US-09-514-247-9	Sequence 9, Appl1
8	21	2.4	8252	1	US-08-046-585-15	Sequence 13, Appl1
9	21	2.4	8252	1	US-08-393-703-15	Sequence 15, Appl1
10	21	2.4	8252	5	PCT-US93-11721-15	Sequence 15, Appl1
11	20	2.3	50	4	US-08-753-247-22	Sequence 22, Appl1
12	20	2.3	51	4	US-08-753-247-23	Sequence 23, Appl1
13	20	2.3	450	4	US-09-370-838-145	Sequence 145, Appl1
14	20	2.3	640	2	US-08-835-099A-16	Sequence 16, Appl1
15	20	2.3	640	3	US-09-157-349-16	Sequence 16, Appl1
16	20	2.3	804	2	US-08-835-099A-10	Sequence 10, Appl1
17	20	2.3	804	3	US-09-157-349-10	Sequence 10, Appl1
18	20	2.3	1028	4	US-08-118-7405-1	Sequence 1, Appl1
19	20	2.3	1028	4	US-08-458-745-1	Sequence 1, Appl1
20	20	2.3	1044	1	US-07-975-526-3	Sequence 3, Appl1
21	20	2.3	1044	4	US-07-974-409C-4	Sequence 425, Appl1
22	20	2.3	1189	1	US-07-781-034-4	Sequence 4, Appl1
23	23	2.3	1189	5	PCT-US92-08328-4	Sequence 4, Appl1
24	20	2.3	1605	4	US-09-124-541-3	Sequence 3, Appl1
25	20	2.3	2130	1	US-07-952-800-1	Sequence 1, Appl1
26	20	2.3	6733	4	US-09-124-541-2	Sequence 2, Appl1
27	20	2.3	10348	2	US-08-457-273B-41	Sequence 41, Appl1

28	20	2.3	10348	3	US-08-556-419-13	Sequence 13, App1
29	20	2.3	10348	4	US-08-556-419-13	Sequence 14, App1
30	20	2.3	10361	1	US-08-246-682A-5	Sequence 5, App11
31	20	2.3	10361	1	US-08-453-265-5	Sequence 5, App11
32	19	2.2	49	1	US-08-155-171B-27	Sequence 27, App1
33	19	2.2	49	1	US-08-155-171B-28	Sequence 28, App1
34	19	2.2	49	2	US-08-435-998-27	Sequence 27, App1
35	19	2.2	49	2	US-08-435-998-28	Sequence 28, App1
36	19	2.2	49	4	US-08-813-781-69	Sequence 69, App1
37	19	2.2	53	1	US-08-155-171B-36	Sequence 36, App1
38	19	2.2	53	2	US-08-435-998-36	Sequence 36, App1
39	19	2.2	81	4	US-09-497-933A-20	Sequence 20, App1
40	19	2.2	91	1	US-08-142-551B-129	Sequence 129, App1
41	19	2.2	91	1	US-08-142-551B-130	Sequence 130, App1
42	19	2.2	276	2	US-08-337-766A-30	Sequence 30, App1
43	19	2.2	440	4	US-09-397-787-328	Sequence 328, App1
44	19	2.2	441	4	US-09-643-597-307	Sequence 307, App1
45	19	2.2	649	4	US-08-998-416-116	Sequence 116, App1

## ALIGNMENTS

```

RESULT 1
US-08-332-766A-1
Sequence 1 Application US/08332766A
Patent No. 5843647
GENERAL INFORMATION:
APPLICANT: JEFFERYS, Alec J.
APPLICANT: ARMOUR, John
TITLE OF INVENTION: SIMPLE TANDEM REPEATS
NUMBER OF SEQUENCES: 125
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
City: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,766A
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9346052.9
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BIRD, Donald J.
REGISTRATION NUMBER: 25,323
REFERENCE/DOCKET NUMBER: 217211/M94/0434/CB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 67146Z CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-332-766A-1

Query Match 2.7%; Score 23; DB 2; Length 377
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 23; Conservative 0; Mismatches 0;

143 TGGTGGTCGATGATGTCGTATGCTG 165

```

DB 123 TGGTGTGATGATGATGATGCTG 145

RESULT 2

US-08-234-783-1  
Sequence 1, Application US/08234783  
Patent No. 5622835  
GENERAL INFORMATION:  
APPLICANT: Herlyn, Meenhard  
APPLICANT: Morris, Jennifer  
APPLICANT: Rauscher III, Frank J.  
APPLICANT: Rodeck, Ulrich  
TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of  
TITLE OF INVENTION: Use Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/234,783  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST48USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 633 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..630

US-08-234-783-1

Query Match 2.4%; Score 21; DB 1; Length 633;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GCACCGCGCGCGCGCGCGCG 46  
DB 208 GCACCGCGCGCGCGCGCGCG 228

RESULT 3

US-08-456-907-1  
Sequence 1, Application US/08456907  
Patent No. 5633142  
GENERAL INFORMATION:  
APPLICANT: Herlyn, Meenhard  
APPLICANT: Morris, Jennifer  
APPLICANT: Rauscher III, Frank J.  
APPLICANT: Rodeck, Ulrich  
TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of  
TITLE OF INVENTION: Use Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,907  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/234,783  
FILING DATE: 28-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST48AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 633 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..630

US-08-456-907-1

Query Match 2.4%; Score 21; DB 1; Length 633;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GCACCGCGCGCGCGCGCGCG 46  
DB 208 GCACCGCGCGCGCGCGCGCG 228

RESULT 4

PCT-US95-05523-1  
Sequence 1, Application PC/TUS9505523  
GENERAL INFORMATION:  
APPLICANT: The Wistar Institute of Anatomy and Biology  
TITLE OF INVENTION: Wt1 Monoclonal Antibodies and  
TITLE OF INVENTION: Methods of Use Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05523  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/234,783  
FILING DATE: 28-APR-1994

```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST48PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; PCT-US95-05523-1

Query Match 2.4%; Score 21; DB 5; Length 633;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 GCACCGCGCGCGCGCGCGCG 46
Db 208 GCACCGCGCGCGCGCGCGCG 228

RESULT 5
US-08-165-315D-3
; Sequence 3, Application US/08165315D
; Patent No. 5525716
; GENERAL INFORMATION:
; APPLICANT: Odd-Arne Olsen
; APPLICANT: Roger Kalla
; TITLE OF INVENTION: Promoter
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: St. Onge, Steward, Johnston & Reens
; STREET: 986 Bedford Street
; CITY: Stamford
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06905
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/165,315D
; FILING DATE: 10 December 1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9324707.0
; FILING DATE: 2 December 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Krinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: 2105-P0001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-324-6155
; TELEFAX: 201-327-1096
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; DESCRIPTION: gene
; FRAGMENT TYPE: gene
; FEATURE:

;
; NAME/KEY: Ltp2 gene
; US-08-165-315D-3

Query Match 2.4%; Score 21; DB 1; Length 1327;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 GTGATGATGATGATGATGATG 168
Db 918 GTGATGATGATGATGATGATG 938

RESULT 6
US-08-264-578-10
; Sequence 10, Application US/08264578
; Patent No. 6391566
; GENERAL INFORMATION:
; APPLICANT: FOLDES, Robert L.
; APPLICANT: ADAMS, Sally-Lin
; APPLICANT: KAMROU, Rajender
; APPLICANT: DUNCAN, H. Scott
; TITLE OF INVENTION: Modulatory Proteins of Human CNS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,578
; FILING DATE: 23-JUN-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/987,953
; FILING DATE: 11-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/261/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1099..3753
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 1099..1152
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1153..3753
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2781..2838
; OTHER INFORMATION: /function= "transmembrane domain"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2895..2958
; OTHER INFORMATION: /function= "transmembrane domain"
```

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2988..3045  
OTHER INFORMATION: /function= "transmembrane domain"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3534..3597  
OTHER INFORMATION: /function= "transmembrane domain"  
US-08-264-578-10

Query Match 2.4%; Score 21; DB 4; Length 4659;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 29 CCGCGCGCGCGCGCGCGCGCG 49  
Db 815 CCGCGCGCGCGCGCGCGCGCG 835

RESULT 7  
US-09-514-247A-9/C  
Sequence 9, Application US/09514247A  
Patent No. 6365361  
GENERAL INFORMATION:  
APPLICANT: TANABE SEIKYAKU CO. LTD.  
APPLICANT: TANIGUCHI, Tomoyasu  
APPLICANT: MIZUKAMI, Junko  
TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA  
FILE REFERENCE: TANIGUCHI-6  
CURRENT APPLICATION NUMBER: US/09/514,247A  
PRIOR FILING DATE: 2000-02-28  
PRIOR APPLICATION NUMBER: PCT/JP98/03734  
PRIOR FILING DATE: 1998-08-24  
PRIOR APPLICATION NUMBER: JP231084/1997  
PRIOR FILING DATE: 1997-08-27  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9  
LENGTH: 8147  
TYPE: DNA  
ORGANISM: human  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (819)..(8147)  
US-09-514-247A-9

Query Match 2.4%; Score 21; DB 4; Length 8147;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 29 CCGCGCGCGCGCGCGCGCGCG 49  
Db 227 CCGCGCGCGCGCGCGCGCGCG 207

RESULT 8  
US-08-046-585-15/C  
Sequence 15, Application US/08046585  
Patent No. 5453362  
GENERAL INFORMATION:  
APPLICANT: Lamarco, Kelly  
APPLICANT: Wilson, Angus  
APPLICANT: Herr, Winship  
TITLE OF INVENTION: A NOVEL EUKARYOTIC TRANSCRIPTION PROTEIN:  
TITLE OF INVENTION: HOST CELL FACTOR  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESS: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/046,585  
FILING DATE: 12-APR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57503-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-046-585-15

Query Match 2.4%; Score 21; DB 1; Length 8252;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 143 TGGTGTGATGATGATGATG 163  
Db 2748 TGGTGTGATGATGATGATG 2728

RESULT 9  
US-08-393-703-15/C  
Sequence 15, Application US/08393703  
Patent No. 5585239  
GENERAL INFORMATION:  
APPLICANT: Lamarco, Kelly  
APPLICANT: Wilson, Angus  
APPLICANT: Herr, Winship  
TITLE OF INVENTION: A NOVEL EUKARYOTIC TRANSCRIPTION PROTEIN:  
TITLE OF INVENTION: HOST CELL FACTOR  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESS: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,703  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57503-2/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8252 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-393-703-15

Query Match 2.4%; Score 21; DB 1; Length 8252;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 TGTGTGATGATGATGATG 163  
|||||  
Db 2748 TGTGTGATGATGATGATG 2728

RESULT 10  
PCT-US93-11721-15/C  
Sequence 15, Application PC/TUS9311721  
GENERAL INFORMATION:  
APPLICANT: Lamarco, Kelly  
APPLICANT: Wilson, Angus  
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:  
TITLE OF INVENTION: HOST CELL FACTOR  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/11721  
FILING DATE: 03-DEC-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: RP-57503-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
PCT-US93-11721-15

Query Match 2.4%; Score 21; DB 5; Length 8252;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 TGTGTGATGATGATGATG 163  
|||||  
Db 2748 TGTGTGATGATGATGATG 2728

RESULT 11  
US-08-753-247-22/C  
Sequence 22, Application US/08753247  
Patent No. 6210929  
GENERAL INFORMATION:  
APPLICANT: SCHLOKAT, Uwe

APPLICANT: FISCHER, Bernhard  
APPLICANT: FALKNER, Falko-Guenther  
APPLICANT: DORNER, Friedrich  
APPLICANT: EIBL, Johann  
TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN  
DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A  
TITLE OF INVENTION: HETEROLOGOUS SEQUENCE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/753,247  
FILING DATE: 22-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AT 1928/95  
FILING DATE: 24-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40433/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-753-247-22

Query Match 2.3%; Score 20; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 TGATGATGATGATGATG 168  
|||||  
Db 42 TGATGATGATGATGATG 23

RESULT 12  
US-08-753-247-23  
Sequence 23, Application US/08753247  
Patent No. 6210929  
GENERAL INFORMATION:  
APPLICANT: SCHLOKAT, Uwe  
APPLICANT: FISCHER, Bernhard  
APPLICANT: FALKNER, Falko-Guenther  
APPLICANT: DORNER, Friedrich  
APPLICANT: EIBL, Johann  
TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN  
DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A  
TITLE OF INVENTION: HETEROLOGOUS SEQUENCE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA

ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/753,247  
FILING DATE: 22-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AT 1928/95  
FILING DATE: 24-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40433/149  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 51 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-753-247-23

Query Match 2.3%; Score 20; DB 4; Length 51;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 149 TGATGATGTGATGTGTG 168  
|||||  
DB 13 TGATGATGTGATGTGTG 32

RESULT 13  
US-09-370-838-145/C  
Sequence 145, Application US/09370838  
Patent No. 6444425  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Mohamath, Radoh  
APPLICANT: Secrist, Heather  
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF  
FILE REFERENCE: 210121.475C1  
CURRENT APPLICATION NUMBER: US/09/370,838  
CURRENT FILING DATE: 1999-08-09  
EARLIER APPLICATION NUMBER: US 09/285,323  
NUMBER OF SEQ ID NOS: 289  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 145  
LENGTH: 450  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-370-838-145

Query Match 2.3%; Score 20; DB 4; Length 450;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 CACCGCGCGCGCGCGCG 46  
|||||  
DB 77 CACCGCGCGCGCGCGCG 58

RESULT 14

US-08-835-099A-16/C  
Sequence 16, Application US/08835099A  
Patent No. 5874277  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasushi  
APPLICANT: NISHI, Kazuori  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/835,099A  
FILING DATE: 04-APR-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 083649/1996  
FILING DATE: 05-APR-1996  
APPLICATION NUMBER: 97105508.2  
FILING DATE: 03-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 47342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 640 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-835-099A-16

Query Match 2.3%; Score 20; DB 2; Length 640;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 CACCGCGCGCGCGCGCG 46  
|||||  
DB 145 CACCGCGCGCGCGCGCG 126

RESULT 15  
US-09-157-349-16/C  
Sequence 16, Application US/09157349  
Patent No. 6068990  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasushi  
APPLICANT: NISHI, Kazuori  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA



COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/157,349  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/835,099  
FILING DATE:  
APPLICATION NUMBER: 97105508.2  
FILING DATE: 03-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 47342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 640 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-157-349-16

Query Match 2.3%; Score 20; DB 3; Length 640;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 CACCGCGCGCGCGCGCGCG 46  
          |||||  
DB 145 CACCGCGCGCGCGCGCGCG 126

Search completed: March 18, 2003, 08:16:26  
Job time : 80 secs

**THIS PAGE BLANK (USPTO)**



[illegible]

```

Sequence 71: Application US/10098841
Publication No. US20020197679A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhimei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Dimauc, Radoje T.
TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
FILE REFERENCE: 784C1P2
CURRENT APPLICATION NUMBER: US/10/098,841
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 71
LENGTH: 1066
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (154)..(867)
US-10-098-841-71

```

QY	542	AGTGGGTGGCGGACACCCCCCAAAACGAACATCTTGCACAGTGCATGTATGGATAGTGGCA	601
Db	545	AGTGGGTGGCGGACACCCCCCAAAACGAACATCTTGCACAGTGCATGTATGGATAGTGGCA	604
QY	602	GGCTGGGGGCGCCCTTGGCCCCCGCAGCATCTGGGCAATCAGCGCCACGTGGTACGGCA	661
Db	605	GGCTGGGGGCGCCCTTGGCCCCCGCAGCATCTGGGCAATCAGCGCCACGTGGTACGGCA	664
QY	662	GGCGGCGGGCGCATGGAGGGGGCGCGGCCCATCTACAGCGAGGTCATCAGGCGCACTACCCGG	721
Db	665	GGCGGCGGGCGCATGGAGGGGGCGCGGCCCATCTACAGCGAGGTCATCAGGCGCACTACCCGG	724
QY	722	GGTCCTCTCTTCCAGCACACACGAGAGCATGGGGCGCCCTCTTGTGTTGGAAGGAGCCCGGC	781
Db	725	GGTCCTCTCTTCCAGCACACACGAGAGCATGGGGCGCCCTCTTGTGTTGGAAGGAGCCCGGC	784
QY	782	TTCACACACACACACATGCGCGCCCTAAAGAGCGCAGCCATCTGGAGCAAAAGAGAGGATA	841
Db	785	TCCACACACACACACATGCGCGCCCTAAAGAGCGCAGCCATCTGGAGCAAAAGAGAGGATA	844
QY	842	AACAGAAAGCACACCTCTC	861
Db	845	AACAGAAAGCACACCTCTC	864

	Query Match	85.9%; Score 740; DB ; length 1066;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	122 TTGTTTCAGATCATCATCTCGTGTGGTGATGATGGTAGTGGTGGTGCACACTGCC	181
Dy		
Dy	125 TTTGTTCCAGATCATCATCTCGTGTGGTGATGATGGTAGTGGTGGTGCACACTGCC	184
QY	182 TGCTGAGCACAATACAGCTGTCTGCACGCTCTTATGAGCCGGACACGCCGGGGCGGA	241
Dy		
Dy	185 TGCTGAGGCACATTACAGCTGTCTGCACGCTCTTATGAGCCGGACACGCCGGGGCGGA	244
QY	242 GGAGAGAAGATGCCCCGTCTCTCAGAGAGATGCTGTGSCCTTGGAGAGACACAGTGTGAG	301
Dy		
Dy	245 GGAAGAGAAGATGCCCCGTCTCTCAGAGAGATGCTGTGSCCTTGGAGAGACACAGTGTGAG	304
QY	302 GCAACGGAATCCAGAGCGCGCAGGCTCTACGGCCCCCTCTGGGCCCAACCGACGCTGGCGCG	361
Dy		
Dy	305 GCACAGGAATCCAGAGCGCGCAGGCTCTACGGCCCCCTCTGGGCCCAACCGACGCTGGCGCG	364
QY	362 TGCCGCCCTTGGCCACAGGGGAGCGCTTCCACGCTTCCAGGCCACCTAATCCGTAACTTGC	421
Dy		
Dy	365 TGCCGCCCTTGGCCACAGGGGAGCGCTTCCACGCTTCCAGGCCACCTAATCCGTAACTTGC	424
QY	422 AGCACGAGATGACCTGCGACCCACACATCTGCTTTCAGACGGGAGAGAGGCCCCACCCT	481
Dy		
Dy	425 AGCACGAGATGACCTGCGACCCACACATCTGCTTTCAGACGGGAGAGAGGCCCCACCCT	484
QY	482 ACCAGGGCCCTTGACACCTTCAGCTTGGGAGACCCGACACAGACTGTGAATCTGAACCGGG	541
Dy		
Dy	485 ACCAGGGCCCTTGACACCTTCAGCTTGGGAGACCCGACACAGACTGTGAATCTGAACCGGG	544

```

1      RESULT 4
2      US-09-796-753-55
3      Sequence 55, Application US/09796753
4      Publication No. US2003027998A1
5      GENERAL INFORMATION:
6      APPLICANT: McCarthy, Sean A.
7      TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
8      FILE REFERENCE: 7853-227-999
9      CURRENT FILING DATE: 2001-03-01
10     PRIOR APPLICATION NUMBER: US/09/796,753
11     PRIOR APPLICATION NUMBER: 09/183,175
12     PRIOR FILING DATE: 1998-10-30
13     PRIOR APPLICATION NUMBER: 09/223,094
14     PRIOR FILING DATE: 1998-12-30
15     PRIOR APPLICATION NUMBER: 09/223,546
16     PRIOR FILING DATE: 1998-12-30
17     PRIOR APPLICATION NUMBER: 09/224,246
18     PRIOR FILING DATE: 1998-12-30
19     PRIOR APPLICATION NUMBER: 09/259,388
20     PRIOR FILING DATE: 1999-02-26
21     PRIOR APPLICATION NUMBER: 60/122,458
22     PRIOR FILING DATE: 1999-03-01
23     PRIOR APPLICATION NUMBER: 09/312,359
24     PRIOR FILING DATE: 1999-05-14
25     PRIOR APPLICATION NUMBER: 09/336,536
26     PRIOR FILING DATE: 1999-06-18
27     PRIOR APPLICATION NUMBER: 09/342,687
28     PRIOR FILING DATE: 1999-06-29
29     PRIOR APPLICATION NUMBER: 09/345,464
30     PRIOR FILING DATE: 1999-06-30
31     PRIOR APPLICATION NUMBER: 09/365,164
32     PRIOR FILING DATE: 1999-07-30
33     PRIOR APPLICATION NUMBER: 09/399,723
34     PRIOR FILING DATE: 1999-09-20
35     PRIOR APPLICATION NUMBER: 09/409,634
36     PRIOR FILING DATE: 1999-09-30
37     PRIOR APPLICATION NUMBER: 09/471,179
38     PRIOR FILING DATE: 1999-12-23
39     PRIOR APPLICATION NUMBER: 09/474,071
40     PRIOR FILING DATE: 1999-12-29
41     PRIOR APPLICATION NUMBER: 09/474,072
42     PRIOR FILING DATE: 1999-12-29
43     PRIOR APPLICATION NUMBER: 09/514,010
44     PRIOR FILING DATE: 2000-02-25
45     PRIOR APPLICATION NUMBER: 09/516,745
46     PRIOR FILING DATE: 2000-03-01
47     PRIOR APPLICATION NUMBER: 09/572,002
48     PRIOR FILING DATE: 2000-05-14

```

Query Match	81.48;	Score 701;	DB 9;	Length 969;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 751; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY .770 AGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGACA 829

RESULT 5  
US-10-000-256A-32  
; Sequence 32, Application US/10000256A  
; Publication No. US20030039983A1

ORGANISM: Homo sapien  
; US-10-000-256A-32

Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1.  $\frac{1}{2}$

Db 1367 TCCTC 1370

## RESULT 6

```
US-09-934-249-14/c
; Sequence 14, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landschulz, Katherine T.
; APPLICANT: Turi, Thomas G.
; APPLICANT: Thompson, John F.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; FILE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (639)...(639)
; OTHER INFORMATION: a, c, g, or t/u
US-09-934-249-14
```

## Query Match

35.1%; Score 302; DB 10; Length 693;  
Best Local Similarity 100.0%; Pred. No. 1.4e-135;

Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 473 CCCACCTTACGAGGCGCCCTGACACCTTCCGAGGAGCCGAGCAGCAGCTGAGAC 532
Db 587 CCCACCTTACGAGGCGCCCTGACACCTTCCGAGGAGCCGAGCAGCAGCTGAGAC 528
QY 533 TGAACCGGAGTGTGCGGCGACCCCAACAGAACCATTTGACAGTAGTACCTGATGG 592
Db 527 TGAACCGGAGTGTGCGGCGACCCCAACAGAACCATTTGACAGTAGTACCTGATGG 468
QY 593 ATATGCGAGCTGTGGGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCC 652
Db 467 ATATGCGAGCTGTGGGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCGCC 408
QY 653 GCTACGCGAGCGGCGGCGCATGAGAGGCGCGCCGACCTACAGCAGGTCATCGGCC 712
Db 407 GCTACGCGAGCGGCGGCGCATGAGAGGCGCGCCGACCTACAGCAGGTCATCGGCC 348
QY 713 ACTACCGGGGTCTCTCTTCAGACAGCAGAGCAGTGGCCGCTCTGCTGAGAG 772
Db 347 ACTACCGGGGTCTCTCTTCAGACAGCAGAGCAGTGGCCGCTCTGCTGAGAG 288
QY 773 GG 774
Db 287 GG 286
```

## RESULT 7

```
US-09-783-590-3464
; Sequence 3464, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruden, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
```

CURRENT APPLICATION NUMBER: US/09/783,590

CURRENT FILING DATE: 2000-02-15

PRIOR APPLICATION NUMBER: 08/420,856

PRIOR FILING DATE: 1995-04-12

PRIOR APPLICATION NUMBER: 08/346,731

PRIOR FILING DATE: 1994-11-21

NUMBER OF SEQ ID NOS: 12485

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3464

LENGTH: 368

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (103)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (225)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (279)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (314)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (349)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (350)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (366)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (367)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: misc feature

LOCATION: (368)

OTHER INFORMATION: n equals a,t,g, or c

US-09-783-590-3464

## Query Match

11.8%; Score 102; DB 10; Length 368;  
Best Local Similarity 100.0%; Pred. No. 1.9e-39;

Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 715 TACCGGGGTCTCTCTTCAGACAGCAGAGCAGTGGCGCCCTCTCTCTGAGGG 774
Db 1 TACCGGGGTCTCTCTTCAGACAGCAGAGCAGTGGCGCCCTCTCTCTGAGGG 60
QY 775 ACCCGGCTCCACACACACATGCGGCCCTAGAGAGCGA 816
Db 61 ACCCGGCTCCACACACACATGCGGCCCTAGAGAGCGA 102
```

## RESULT 8

```
US-09-783-590-3488
; Sequence 3488, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruden, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
```

PRIOR APPLICATION NUMBER: 09/223,546

```

; APPLICANT: Lee, Richard T.
; APPLICANT: Landschulz, Katherine T

```



RESULT 12  
US-09-864-761-20542  
Sequence 20542, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aomico-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21

```

RESULT 13
US-09-864-761-3776
; Sequence 3776, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

```

```

PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/2234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3776
LENGTH: 1964
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005291.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.7
US-09-864-761-3776
```

```
Query Match
Best Local Similarity 3.7%; Score 32; DB 10; Length 1964;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

DB 402 GTGGTGTGATGATGATGATGATGATGATGATGAT 433
142 GTGGTGTGATGATGATGATGATGATGATGATGAT 173
```

```

RESULT 14
US-09-864-761-3936
Sequence 3936, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263,6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
```

```

PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3936
LENGTH: 401
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007249.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7
US-09-864-761-3936
```

```
Query Match
Best Local Similarity 3.4%; Score 29; DB 10; Length 401;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

DB 264 TGGTGTGATGATGATGATGATGATGATGATGAT 292
143 TGGTGTGATGATGATGATGATGATGATGATGAT 171
```

```

RESULT 15
US-09-864-761-20699
Sequence 20699, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263,6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
```

```
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 20699
? LENGTH: 446
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC007249.2
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 10
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 14
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 13
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.7
? OTHER INFORMATION: NT HIT: AL161539.2, EVALUATE 3.70e+00
US-09-864-761-20699

Query Match 3.48; Score 29; DB 10; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 143 TGGTGTGATGATGTGATGTGTCGTCG 171
Db 143 TGGTGTGATGATGTGATGTGTCGTCG 171
```

Search completed: March 18, 2003, 08:17:36  
Job time : 82 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 07:15:25 ; Search time 1293 Seconds  
(without alignments)  
10784.466 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861  
Sequence: 1 atgacgcgtctgtatgagggtc.....aacagaagaagacaccctctc 861

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estma:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_hum:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtc:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_fod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	614	71.3	967	14	B0641849
2	566	65.7	1046	14	BM922276
3	538	64.8	609	14	B0636742
4	550	63.9	551	13	BM141979
5	499	58.0	916	14	B0954555
6	497	57.7	890	14	B0690750

7	482	56.0	805	9	AL558881	AL558881
8	470	54.6	729	14	B0575741	B0575741
9	420	49.2	949	9	AL517150	AL517150
10	420	48.8	655	14	B0691705	B0691705
11	420	48.8	1280	14	B0691500	B0691500
12	417	48.4	782	14	B0015170	B0015170
13	417	48.4	898	9	AL558882	AL558882
14	393	45.6	461	14	BM712680	BM712680
15	389	45.2	633	14	BM714472	BM714472
16	374	43.4	844	14	B0686793	B0686793
17	369	42.9	559	12	BE855409	BE855409
18	353	41.0	730	14	BM677602	BM677602
19	351	40.8	626	14	BM974296	BM974296
20	335	38.9	570	14	B0575582	B0575582
21	330	38.3	552	14	BM713900	BM713900
22	302	35.1	693	9	AL761441	AL761441
23	279	32.4	547	14	BM676516	BM676516
24	276	32.1	613	12	BG680325	BG680325
25	257	29.8	308	14	BM705514	BM705514
26	243	28.2	502	9	AL921394	AL921394
27	238	27.6	446	14	BM681946	BM681946
28	237	27.5	453	9	AA917446	AA917446
29	235	27.3	437	9	AL936228	AL936228
30	217	25.2	464	12	BF026695	BF026695
31	206	23.9	518	9	AL885001	AL885001
32	198	23.0	451	9	AL493698	AL493698
33	198	23.0	990	12	BG675643	BG675643
34	195	22.6	404	9	AL925027	AL925027
35	194	22.5	634	9	AL826012	AL826012
36	191	22.2	588	9	AL377498	AL377498
37	191	22.2	619	9	AL742327	AL742327
38	187	21.7	380	10	BE138909	BE138909
39	183	21.3	508	9	AA595115	AA595115
40	182	21.1	284	10	AA452945	AA452945
41	179	20.8	467	10	AM007283	AM007283
42	177	20.6	397	12	BF446904	BF446904
43	177	20.6	430	10	AW204238	AW204238
44	177	20.6	463	12	BF939262	BF939262
45	177	20.6	674	9	AL972096	AL972096

## ALIGNMENTS

RESULT 1  
B0641849  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

B0641849  
AGENCOURT.8287174 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6292265  
B0641849  
B0641849.1 GI:21766021  
EST.

Human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L10CM2493 row: 9 column: 18  
High quality sequence stop: 571.  
Location/Qualifiers

FEATURES  
source

1. 967

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6292265"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pORF7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald W. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC library. |"

BASE COUNT      194 a       334 c       296 g       143 t
ORIGIN

Query Match          71.3%; Score 614; DB 14; Length 967;
Best Local Similarity 99.7%; Prid. No. 8.3e-259;
Matches 714; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY   112 GACCTGGAGTTTGTCATCATCATCATGTGTGTATGATGATGGTAGTGTTGTGTG 171
DB    1 GACTGGAGTTTGTGATGATCATCATCATGTGTGTATGATGATGGTAGTGTTGTGTG 60

QY   172 ATACGGTCCGTGTGAGCCCTACAAGCTGTCTGCACGGTCTCTTCATCAGCCGGCACAGC 231
DB    61 ATACGTTCCGTGTGAGCCCTACAAGCTGTCTGCACGGTCTCTTCATCAGCCGGCACAGC 120

QY   232 CAGGGCGGAGAGAAGATGCCCTGTCTCTGAAGAGATGCTTGAGGCCCTGSGAGAG 291
DB    121 CAGGGCGGAGAGAAGATGCCCTGTCTCTGAAGAGATGCTTGAGGCCCTGSGAGAG 180

QY   292 ACAGTTCAGGCAACGGAATCCCAGAGCCGACAGTTACGCCCCGCTCGGCCACCGAC 351
DB    181 ACAGTTCAGGCAACGGAATCCCAGAGCCGACAGTTACGCCCCGCTCGGCCACCGAC 240

QY   352 CGCGTGGCGGTGCGCCCTTTCGCCCGAGGGAGCGCTTACCAGCCCTTCAGCCACTAT 411
DB    241 CGCGTGGCGGTGCGCCCTTTCGCCCGAGGGAGCGCTTACCAGCCCTTCAGCCACTAT 300

QY   412 CCGTACTGCACACAGATGCAGCTGCCACCCACCATCTGCTGTACAGCGGGAGAG 471
DB    301 CCGTACTGCACACAGATGCAGCTGCCACCCACCATCTGCTGTACAGCGGGAGAG 360

QY   472 CCCCAACCTTACAGGGCCCCGACGCCCTTCCAGCTTCCGGGAGACCCGAGCAGCTGGAA 531
DB    361 CCCCAACCTTACAGGGCCCCGAGCCCTTCCAGCTTCCGGGAGACCCGAGCAGCTGGAA 420

QY   532 CTGAACCGGAGCTCGTGTGCGCGCACCCCCAACAACCATTTGTGACAGTAGCTGATG 591
DB    421 CTGAACCGGAGCTCGTGTGCGCGCACCCCCAACAACCATTTGTGACAGTAGCTGATG 480

QY   592 GATTGTGCAAGCTGGGGGGGCCCCCTGCCCCCAGCACTAACTGGGGCATCAGCGCCACG 651
DB    481 GATTGTGCAAGCTGGGGGGGCCCCCTGCCCCCAGCACTAACTGGGGCATCAGCGCCACG 540

QY   652 TGTACAGCGCAGCGGGCGCATGGAGGGGGCGGCCACCACTTACAGAGAGTAGTGGC 711
DB    541 TGTACAGCGCAGCGGGGGCGTATGGAGGGGGCGGCCACCACTTACAGAGAGTAGTGGC 600

QY   712 CACTAACCGGGGCTCTCTTCCAGCACCCAGAGAGACAGTGGGGCGCCCTCTTGTGGAG 771
DB    601 CACTAACCGGGGCTCTCTTCCAGCACCCAGAGAGAGTGGGGCGCCCTCTTGTGGAG 660

QY   772 GGGAGCCGGGCTCCACACACACACATGGCGGCCCTTAAGAGAGCAGCCATTTGAG 827
DB    661 GGGAGCCGGGCTCCACACACACACATGGCGGCCCTTAAGAGAGCAGCCATTTGAG 716

RESULT 2
BM922276
JOCUS

```

DEFINITION	AGENCOURT_6707077 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754437		
ACCESSION	5', mRNA sequence.		
VERSION	BM922276		
KEYWORDS	BM922276.1 GI:19372655		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1046)		
JOURNAL	NIH_MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: L14M12791 row: n column: 06 High quality sequence stop: 671.		
FEATURES	Location/Qualifiers		
Source	1..1046		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5754437"		
	/clone_11b="NIH_MGC_115"		
	/lab_host="DH10B"		
	/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 021. Note: this is a NIH_MGC Library."		
BASE COUNT	205 a 391 c 298 g 150 t 2 others		
ORIGIN			
Query Match	65.7%; Score 566; DB 14; Length 1046;		
Best Local Similarity	99.7%; Pred. No. 9.6e-238;		
Matches	666; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
0Y	1 ATGACCGCGTTGATGGGGGTCACAGACCGCGCGCCGCGCGGAGCGCCATATGTC 60		
Db			
67	ATGACCGCGTTGATGGGGGTCACAGACCGCGCGCGCGCGGAGCGCCATATGTC 126		
0Y	61 TCCTGCACGTCACATGCAAGCGCTCTTTGTTCAGAGCATGAGATCACGGAGCTGAG 120		
Db			
127	TCCTGCACGTCACATGCAAGCGCTCTTTGTTCAGAGCATGAGATCACGGAGCTGAG 186		
0Y	121 TTTGTTCAGATCATCATCATCTGCTGCTGCTGATGATGATGATGATGATGATGATG 180		
Db			
187	TTTGTTCAGATCATCATCATCTGCTGCTGCTGATGATGATGATGATGATGATGATG 246		
0Y	181 CTGCTGACCCATCAAGAGCTGTCTGCAAGGCTCTTATCATAGCCGGACAGCCAGGGGGCG 240		
Db			
247	CTGCTGACCCATCAAGAGCTGTCTGCAAGGCTCTTATCATAGCCGGACAGCCAGGGGGCG 306		
0Y	241 AGGAGAAAGATGCGCGTCTCAGAAAGATGCTGGCGCTCGGAGAGACACATGTCA 300		
Db			
307	AGGAGAAAGATGCGCGTCTCAGAAAGATGCTGGCGCTCGGAGAGACACATGTCA 366		
0Y	301 GGCAGAGGAATCCCAAGAGCGCGAGGCTCAAGCGCGCGCTCGGCGACAGCAGCGCTGGCC 360		
Db			
367	GGCAGAGGAATCCCAAGAGCGCGAGGCTCAAGCGCGCGCTCGGCGACAGCAGCGCTGGCC 426		
0Y	361 GTGGCGCGCTTCGCCAGGGGAGCGCTTCACCGCTTCACAGCCACCATATTCGACTCG 420		

```

|||||
Db 427 GTGCGCCCTTGGCCAGGGGAGCGGTTCCACGGCTTCCAGGCCACCTTCCGTAACCTG 486
QY 421 CAGCAGAGATGACCTGCCACCCACCACATCTGCTGTGACAGCGGGAGAGACCCCAACC 480
Db 487 CACACAGAGATGACCTGCCAGCGCCACCATCTGCTGTGACAGCGGGAGAGACCCCAACC 546
QY 481 TACAGAGGCGCCCTGCACCCCTGCAGCTTCCGGAGCCCGAGAGAGAGTGAACGAACCGG 540
Db 547 TACAGAGGCGCCCTGCACCCCTGCAGCTTCCGGAGCCCGAGAGAGAGTGAACGAACCGG 606
QY 541 GACTCGGTGGCGGACCCCAACAGAACATCTGACAGATGACCTGATGATGATGATGATGATG 600
Db 607 GACTCGGTGGCGGACCCCAACAGAACATCTGACAGATGACCTGATGATGATGATGATGATG 666
QY 601 AGGCTGGGGGGGCGCTGCCCCCAGACAGTACTGGGGGATCAGCGCCACGTCTACGGC 660
Db 667 AGGCTGGGGGGGCGCTGCCCCCAGACAGTACTGGGGGATCAGCGCCACGTCTACGGC 726
QY 661 AGGGGGG 668
Db 727 AGCGGCGG 734

RESULT 3
B0636742 609 bp mRNA linear EST 15-JUL-2002
LOCUS hd13h06.y1 Human Retina cDNA (un-normalized, unamplified): hd/he
DEFINITION Homo sapiens cDNA clone hd13h06 5', mRNA sequence.
ACCESSION B0636742
VERSION B0636742.1 GI:21761201
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 609)
AUTHORS Wislow,G., Bernstein,S.L., Wyatt,M.K., Ray,S., Behal,A., Touchman
J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of human retina for the NEIRBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
JOURNAL Contact: Wislow G
COMMENT Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 13 row: h column: 06
Seq primer: M13Rpl reverse primer (ABI).
FEATURES
source
1. 609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="hd13h06"
/clone_lib="Human Retina cDNA (un-normalized, unamplified)
): hd/he"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10b"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the Superscript Plasmid System full
details of which are contained in the manufacturer's
instruction manual (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTAGTCTAGATCGGAGCGCGCC(T)15-3'

```

```

BASE COUNT 114 a 238 c 182 g 75 t
ORIGIN
Query Match 64.8%; Score 558; DB 14; Length 609;
Best local Similarity 99.8%; Pred. No. 3.1e-234;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 219 CAGCCGGCAGAGCCAGGCGGAGAGAGAAGATGCCCTTCTCCAGAGAGATGCCGTG 278
Db 1 CAGCCGGCAGAGCCAGGCGGAGAGAGAAGATGCCCTTCTCCAGAGAGATGCCGTG 60
QY 279 GCCCTGGGAAGACAGATGTGACGCAACGGAATCCAGAGCCGAGGTCTACGCCGCC 338
Db 61 GCCCTGGGAAGACAGATGTGACGCAACGGAATCCAGAGCCGAGGTCTACGCCGCC 120
QY 339 TCGGCCACGAGCGGCTGACCGTGGCCGCTTCCGCCACGCGGAGGCGTTCACGCGCTT 398
Db 121 TCGGCCACGAGCGGCTGACCGTGGCCGCTTCCGCCACGCGGAGGCGTTCACGCGCTT 180
QY 399 CCAAGCCACCTATCCGTACCTGACAGACAGATGACCTGCAACCCACATCTGCTGTG 458
Db 181 CCAAGCCACCTATCCGTACCTGACAGACAGATGACCTGCGGCCACATCTGCTGTG 240
QY 459 AGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
Db 241 AGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 519 GCAGAGAGCTGGAAGTGAACCGGGAGTGGTGGCGGACCCCAACAGAACATCTTGA 578
Db 301 GCAGAGAGCTGGAAGTGAACCGGGAGTGGTGGCGGACCCCAACAGAACATCTTGA 360
QY 579 CAGTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 638
Db 361 CAGTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 639 CATTACGCGCCACCTGCTACAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 698
Db 421 CATTACGCGCCACCTGCTACAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 699 CGAGGTCATCGGCGCAGTACCGCGGGGTCCTCTCCAGACAGACAGAGAGAGAGAGAG 758
Db 481 CGAGGTCATCGGCGCAGTACCGCGGGGTCCTCTCCAGACAGAGAGAGAGAGAGAGAG 540
QY 759 CTCCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818
Db 541 CTCCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 819 CATCTGGAG 827
Db 601 CATCTGGAG 609

RESULT 4
BM141979 551 bp mRNA linear EST 12-MAR-2002
LOCUS BM141979
DEFINITION I25b11.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5677341 5' similar to TR:09UDJ3 09UDJ3 DJ718J7.1
//, mRNA sequence.
ACCESSION BM141979
VERSION BM141979.1 GI:117152046
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Seacrest,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Merra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R., Williams,T.

```





[illegible]

Query Match	Best Local Similarity	Score	DB	Length
Matches 5477, Conservative	99.8%	Pred. No. 1.8e-207	Mismatches 1	Indels 0
			Gaps 0	
110	CGGAGCTGGAGTTGTTCAGATCATCATCTGTTGGTGTATGTATGTGTGTG	169		
128	CGGAGCTGGAGTTGTTCAGATCATCATCTGTTGGTGTATGTATGTGTGTG	187		
170	TGATACGTGCTGCTGAGCCACTTCAAGCTGTCTGACGGTCTTTCATCAGCCGGCACA	229		
188	TGATACGTGCTGCTGAGCCACTTCAAGCTGTCTGACGGTCTTTCATCAGCCGGCACA	247		
230	GCCAGGGGGGAGAGAAAGTGGCTGTCTCGAAGAGATGCTGTGGCCCTCGGAGAGA	289		
248	GCCAGGGGGGAGAGAAAGTGGCTGTCTCGAAGAGATGCTGTGGCCCTCGGAGAGA	307		
290	GCACAGTGTCAAGGCAAGGAAATCCAGAGCCGACAGGTCTACGCCGCCCTCGGCCACCG	349		
308	GCACAGTGTCAAGGCAAGGAAATCCAGAGCCGACAGGTCTACGCCGCCCTCGGCCACCG	367		
350	ACCGCTGGCGCTGCGCGCCCTTCGCCCAAGCGGAGAGCGCTTCCAGCGCTTCCAGCCCACT	409		
368	ACCGCTGGCGCTGCGCGCCCTTCGCCCAAGCGGAGAGCGCTTCCAGCGCTTCCAGCCCACT	427		
410	ATTCGTACCTGTGAGACGATCGACTGCGCCACCCACCATCTCGCTGTCTGACAGCGGGAGG	469		
428	ATTCGTACCTGTGAGACGATCGACTGCGCCACCCACCATCTCGCTGTCTGACAGCGGGAGG	487		
470	AGCCCCCAACCTTACAGAGGGGCCCTCGACCCCTTCAGCTTTCGGGAGCCCGAGACGAGCTGG	529		
488	AGCCCCCAACCTTACAGAGGGGCCCTCGACCCCTTCAGCTTTCGGGAGCCCGAGAGCTGG	547		
530	AACTAACAACCGGAGTGTGCGCGCAACCCCAACAGAACATCTTTCGACAGTACCTGA	589		
548	AACTAACAACCGGAGTGTGCGCGCAACCCCAACAGAACATCTTTCGACAGTACCTGA	607		
590	TGGATAGTGCAGGCTGGGGGGGCCCTCGCCCGCCAGAGTAACTCGGGCATCAGCGCA	649		
608	TGGATAGTGCAGGCTGGGGGGGCCCTCGCCCGCCAGAGTAACTCGGGCATCAGCGCA	667		
650	GCTGCTAC	657		
668	CGTGTAC	675		

BASE COUNT 168 a 303 c 273 g 146 t

ORIGIN

Query Match 57.7% Score 497: DB 14: Length 890:

Best Local Similarity 99.8% Pred. No. 1.8e-207:

Matches 5477, Conservative 0: Mismatches 1: Indels 0: Gaps 0:

110 CGGAGCTGGAGTTGTTCAGATCATCATCTGTTGGTGTATGTATGTGTGTG 169

128 CGGAGCTGGAGTTGTTCAGATCATCATCTGTTGGTGTATGTATGTGTGTG 187

170 TGATACGTGCTGCTGAGCCACTTCAAGCTGTCTGACGGTCTTTCATCAGCCGGCACA 229

188 TGATACGTGCTGCTGAGCCACTTCAAGCTGTCTGACGGTCTTTCATCAGCCGGCACA 247

230 GCCAGGGGGGAGAGAAAGTGGCTGTCTCGAAGAGATGCTGTGGCCCTCGGAGAGA 289

248 GCCAGGGGGGAGAGAAAGTGGCTGTCTCGAAGAGATGCTGTGGCCCTCGGAGAGA 307

290 GCACAGTGTCAAGGCAAGGAAATCCAGAGCCGACAGGTCTACGCCGCCCTCGGCCACCG 349

308 GCACAGTGTCAAGGCAAGGAAATCCAGAGCCGACAGGTCTACGCCGCCCTCGGCCACCG 367

350 ACCGCTGGCGCTGCGCGCCCTTCGCCCAAGCGGAGAGCGCTTCCAGCGCTTCCAGCCCACT 409

368 ACCGCTGGCGCTGCGCGCCCTTCGCCCAAGCGGAGAGCGCTTCCAGCGCTTCCAGCCCACT 427

410 ATTCGTACCTGTGAGACGATCGACTGCGCCACCCACCATCTCGCTGTCTGACAGCGGGAGG 469

428 ATTCGTACCTGTGAGACGATCGACTGCGCCACCCACCATCTCGCTGTCTGACAGCGGGAGG 487

470 AGCCCCCAACCTTACAGAGGGGCCCTCGACCCCTTCAGCTTTCGGGAGCCCGAGACGAGCTGG 529

488 AGCCCCCAACCTTACAGAGGGGCCCTCGACCCCTTCAGCTTTCGGGAGCCCGAGAGCTGG 547

530 AACTAACAACCGGAGTGTGCGCGCAACCCCAACAGAACATCTTTCGACAGTACCTGA 589

548 AACTAACAACCGGAGTGTGCGCGCAACCCCAACAGAACATCTTTCGACAGTACCTGA 607

590 TGGATAGTGCAGGCTGGGGGGGCCCTCGCCCGCCAGAGTAACTCGGGCATCAGCGCA 649

608 TGGATAGTGCAGGCTGGGGGGGCCCTCGCCCGCCAGAGTAACTCGGGCATCAGCGCA 667

650 GCTGCTAC 657

668 CGTGTAC 675

RESULT 7

AL558881/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Euarchyotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

805 bp mRNA Linear EST 16-FEB-2001

AL558881.LT1.NFL008.TC2 Homo sapiens cdna clone GS0D015YF12 3

prime, mRNA sequence.

AL558881

AL558881.1 GI:12903836

EST.

REFERENCE 1 (bases 1 to 805)  
 AUTHORS Li W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

## FEATURES

location/Qualifiers

1..805

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CSDBJ015YF12"

/clone\_1db="UTL\_NFL008\_TC2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax: (1) 301 610 8371

Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com"

BASE COUNT 106 a 248 c 289 g 161 t 1 others

## ORIGIN

Query Match 56.0%; Score 482; DB 9; Length 805;

Best Local Similarity 99.5%; Pred. No. 6.9e-201; Mismatches 3; Indels 0; Gaps 0;

Matches 632; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

227 ACAGCCAGGCGGAGAGAGAGATGCCCTGTCTCAGAGAGATCCCTGTGCGG 286  
 |||||||  
 755 ACAGCCAGGCGGAGAGAGATGCCCTGTCTCAGAGAGATCCCTGTGCGG 696  
 |||||||  
 287 AGAGCAGATGTCAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 346  
 |||||||  
 695 AGAGCAGATGTCAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 636  
 |||||||  
 347 CGAGCCGCTGGCGCTGCGCCCTTCCAGCGGAGCGCTTCCAGCGCTTCCAG 406  
 |||||||  
 635 CGAGCCGCTGGCGCTGCGCCCTTCCAGCGGAGCGCTTCCAGCGCTTCCAG 576  
 |||||||  
 407 CCTATCCGATCTGACAGCAGAGATGACCTGCGCCACCTTCTGCTTACAG 466  
 |||||||  
 575 CCTATCCGATCTGACAGCAGAGATGACCTGCGCCACCTTCTGCTTACAG 516  
 |||||||  
 467 AGAGCCCGCCACCTTCCAGCGGAGCGCTTCCAGCGGAGCGCTTCCAGCG 526  
 |||||||  
 515 AGAGCCCGCCACCTTCCAGCGGAGCGCTTCCAGCGGAGCGCTTCCAGCG 456  
 |||||||  
 527 TGGAACTGAACCGGAGTGGTGGCGCACCCCAAGAACCATCTTGCAGTGA 586  
 |||||||  
 455 TGGAACTGAACCGGAGTGGTGGCGCACCCCAAGAACCATCTTGCAGTGA 396  
 |||||||  
 587 TGAATGATGTCAGAGCGTGGGCGGCGCTGCGCCCGCCAGCATTAATCG 646  
 |||||||  
 395 TGAATGATGTCAGAGCGTGGGCGGCGCTGCGCCCGCCAGCATTAATCG 336  
 |||||||  
 647 CCAGTGTACGAGCGGAGCGGCGATGAGAGGCGCGCCACCTACAGGAGTCA 706  
 |||||||  
 335 CCAGTGTACGAGCGGAGCGGCGATGAGAGGCGCGCCACCTACAGGAGTCA 276  
 |||||||  
 707 TCGGCACTACCGCGGCTCTCTTCAGACAGAGAGAGAGAGAGAGAGAGAG 766  
 |||||||  
 275 TCGGCACTACCGCGGCTCTCTTCAGACAGAGAGAGAGAGAGAGAGAGAG 216  
 |||||||  
 767 TGGAGGAGAGCGGCTTCCAGCAGACACATCGCGCCCTAGAGAGAGAGCAT 826  
 |||||||  
 215 TGGAGGAGAGCGGCTTCCAGCAGACACATCGCGCCCTAGAGAGAGAGCAT 156  
 |||||||

QY 827 GCAAAGAGAGATTAACAGAGAGACACCTCTC 861  
 |||||||  
 DB 155 GCAAAGAGAGATTAACAGAGAGACACCTCTC 121

## RESULT 8

BO575741/c

LOCUS

DEFINITION

UT-H-E21-bbg-h-14-0-UI.s1 NCL CGAP\_Ch2 Homo sapiens

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

QY 341 GGCCACGACCGCGCTGCGCGCTTCCAGCGGAGCGCTTCCAGCGCTTCC 400  
 |||||||  
 DB 729 GGCCACGACCGCGCTGCGCGCTTCCAGCGGAGCGCGCTTCCAGCGCTTCC 670  
 |||||||  
 QY 401 AGCCACCTATCCGATCTGACAGCAGAGATGACCTGCCACCCACCATCTG 460  
 |||||||  
 DB 669 AGCCACCTATCCGATCTGACAGCAGAGATGACCTGCCACCCACCATCTG 610  
 |||||||

## FEATURES

source

location/Qualifiers

1..729

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-H-E21-bbg-h-14-0-UI"

/clone\_1db="NCL CGAP\_Ch2"

/tissue\_type="Chondrosarcoma Grade II"

/dev\_stage="Adult"

/lab\_host="DHI0B (Life Technologies)"

/note="Organ: Left Pelvis; Vector: pRTT3-Pac (Pharmacia)

with a modified polylinker; Site1: EcoR I; Site2: Not I;

NCL CGAP\_Ch2 is a normalized cDNA library containing the

following tissue(s): Chondrosarcoma Grade II. The library

was constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pRTT3-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

(dTT18 tail). The sequence tag for this library is

TGATCAGCCT.

TAG\_LIB-UI-H-E21

TAG\_TISSUE=grade-2-chondrosarcoma

TAG\_SEQ=ATCTAATATG"

BASE COUNT 101 a 213 c 253 g 162 t

## ORIGIN

Query Match 54.6%; Score 470; DB 14; Length 729;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-195; Mismatches 1; Indels 0; Gaps 0;





SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
JOURNAL 1 (bases 1 to 782)  
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapds-femail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
source  
1..782  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5834635"  
/clone\_lib="NCI\_CGAP\_EPI"  
/tissue\_type="Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Left Pubic Bone; Vector: pT73-Pac  
(Pharmacia) with a modified polylinker; Site 1: EcoR I;  
Site 2: Not I; NCI CGAP EPI is a normalized cDNA library  
containing the following tissue(s): Chondrosarcoma cell  
line C5. The library was constructed according to Bonaldo  
, Lennon and Soares, Genome Research, 6:791-806, 1996.  
First strand cDNA synthesis was primed with an oligo-dT  
primer containing a Not I site. Double stranded cDNA was  
ligated to an EcoR I adaptor, digested with Not I, and  
cloned directionally into pT73-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is GCTCAAGGCT.  
TAG\_LIB=UI-H-ED1  
TAG\_TISSUE=chondrosarcoma  
TAG\_SEQ=CCTCAAGGCT"

BASE COUNT 109 a 223 c 271 g 176 t 3 others  
ORIGIN

Query Match 48.4%; Score 417; DB 14; Length 782;  
Best local Similarity 99.6%; Pred. No. 2.3e-172;  
Matches 517; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 343 CCCACGACGCGCTGGCGGCGCCCTTCGCCAGCGGAGCGCTTCACCGCTCCAG 402  
|||||  
Db 725 CCCACGACGCGCTGGCGGCGCCCTTCGCCAGCGGAGCGCTTCACCGCTCCAG 666  
|||||

QY 403 CCCACCTATCCGTACCTGCACAGAGATGCAGCTGCACACCACCATCTGCTGACAG 462  
|||||  
Db 665 CCCACCTATCCGTACCTGCACAGAGATGCAGCTGCACACCACCATCTGCTGACAG 606  
|||||

QY 463 GGGAGAGAGCCCACTACAGAGGCGCCCTGCACCCCTCCAGCTTGGGAGCCGAGAG 522  
|||||  
Db 605 GGNAGAGAGCCCACTACAGAGGCGCCCTGCACCCCTCCAGCTTGGGAGCCGAGAG 546  
|||||

QY 523 CAGCTGAATGAACCGGAGTGGTGGCGGACCCCAACAGAACCATCTTGACAGT 582  
|||||  
Db 545 CAGCTGAATGAACCGGAGTGGTGGCGGACCCCAACAGAACCATCTTGACAGT 486  
|||||

QY 583 GACCTGATGATGATGGCAGGCTGGGGGCGCCCGCCCGACAGAAATTCGGGAGTC 642  
|||||  
Db 485 GACCTGATGATGATGGCAGGCTGGGGGCGCCCGCCCGACAGAAATTCGGGAGTC 426  
|||||

QY 643 AGCGCCAGTGTACTAGGCGAGGGGCGCATGTAGAGGGGCGCGCCCACTACAGCGAG 702  
|||||

|||||  
Db 425 AGCGCCAGTGTACTAGGCGAGGGGCGCATGTAGAGGGGCGCGCCCACTACAGCGAG 366  
|||||

QY 703 GTCATCGGCACATACCCGGGCTCTCTTCCAGACACAGAGAGATGGGCGCCCTCC 762  
|||||

Db 365 GTCATCGGCACATACCCGGGCTCTCTTCCAGACACAGAGAGATGGGCGCCCTCC 306  
|||||

QY 763 TTGCTGGAGGGGACCGGCTCCACACACATCGGCGCCCTTGAAGAGGCGACCATC 822  
|||||

Db 305 TTGCTGGAGGGGACCGGCTCCACACACATCGGCGCCCTTGAAGAGGCGACCATC 246  
|||||

QY 823 TGGAGCAAGAGAGATTAACAGAAAGACACCTCTC 861  
|||||

Db 245 TGGAGCAAGAGAGATTAACAGAAAGACACCTCTC 207  
|||||

RESULT 13  
AL558882 898 bp mRNA linear EST 16-FEB-2001  
LOCUS AL558882 LTI\_NFL008.TC2 Homo sapiens cDNA clone CS0D015YF12 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL558882  
VERSION AL558882.1 GI:12903838  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1. (bases 1 to 898)  
AUTHORS Li, W.-B., Gruber, C., Jesssee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequences  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source  
1..898  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0D015YF12"  
/clone\_lib="LTI\_NFL008\_TC2"  
/sex="male"  
/tissue\_type="T cells from T cell leukemia"  
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
Life Technologies. Contact : Feng Liang Life Technologies,  
a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 196 a 310 c 260 g 126 t 6 others  
ORIGIN

Query Match 48.4%; Score 417; DB 9; Length 898;  
Best local Similarity 99.5%; Pred. No. 2.3e-172;  
Matches 637; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 177 GTGCTGTGAGCCACTCAAGCTGTGCAAGGCTTCATCAAGCGGACAGCAGAGG 236  
|||||

Db 1 GTGCTGTGAGCCACTCAAGCTGTGCAAGGCTTCATCAAGCGGACAGCAGAGG 60  
|||||

QY 237 GCGAGAGAGAGATGCTCTCTCAGAGGATGCTTGCGCTCGAGAGACAGT 296  
|||||

Db 61 GCGAGAGAGAGATGCTCTCTCAGAGGATGCTTGCGCTCGAGAGACAGT 120  
|||||

QY 297 GTCAAGCAAGGAATCCAGAGCGGAGGTACGCGCCCGCTGGGCCACGACGCGCT 356  
|||||

Db 121 GTCAAGCAAGGAATCCAGAGCGGAGGTACGCGCCCGCTGGGCCACGACGCGCT 179  
|||||

QY 357 GGGCTGCGCCCTTGGCCGAGGAGGCGCTTCCACCGCTTCAGCCCACTATCCGTA 416  
 DB 180 GGGCGTGGCCCTTGGCCGAGGAGGCGCTTCCACCGCTTCAGCCCACTATCCGTA 239  
 QY 417 CCGGAGCAGAGATGAGCTGGCCACCAACATCTGCTGTGAGAGGGAGAGCCCGC 476  
 DB 240 CCGGAGCAGAGATGAGCTGGCCACCAACATCTGCTGTGAGAGGGAGAGCCCGC 239  
 QY 477 ACCCTACAGAGGCGCTTCAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 536  
 DB 300 ACCCTACAGAGGCGCTTCAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 359  
 QY 537 CCGGAGTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 596  
 DB 360 CCGGAGTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419  
 QY 597 TGGCAGGCTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 656  
 DB 420 TGGCAGGCTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 479  
 QY 657 CCGCAGGAGGAG 716  
 DB 480 CCGCAGGAGGAG 539  
 QY 717 CCGGAGGCTGCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 776  
 DB 540 CCGGAGGCTGCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 559  
 QY 777 CCGGCTCCACACACACACATGCGGCGCCCTAGAGAGCGCA 816  
 DB 600 CCGGCTCCACACACACATGCGGCGCCCTAGAGAGCGCA 639

RESULT 14  
 LOCUS BM712680 461 bp mRNA linear EST 28-FEB-2002  
 DEFINITION UI-E-EJ0-ahg-a-10-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone  
 ACCESSION BM712680  
 VERSION UI-E-EJ0-ahg-a-10-0-UI 5', mRNA sequence.  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 461)  
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msouares@blue.weeg.uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.

FEATURES  
 source Location/Qualifiers  
 1..461  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-EJ0-ahg-a-10-0-UI"  
 /clone\_1b="UI-E-EJ0"  
 /tissue\_type="fetal eyes, lens, eye anterior segment,  
 optic nerve, retina, Retina foveal and Macular, RPE and

"choroid"  
 /dev\_stage="fetal and adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: eye; Vector: p773-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 UI-E-EJ0 is a subtracted cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into p773-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are: fetal eyes, AGAATCAGA  
 ; lens, CGATTACGCA; eye anterior segment, ATGCCGAT;  
 optic nerve, CCATTAAATG; retina, CCGCC; Retina foveal and  
 Macular, GTCC; RPE and Choroid, ACCTA. This library was  
 created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."

BASE COUNT 97 a 166 c 141 g 57 t  
 ORIGIN  
 Query Match 45.6%; Score 393; DB 14; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-162;  
 Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 GAGCCCCCAGCCCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 528  
 DB 1 GAGCCCCCAGCCCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60  
 QY 529 GAACGAGAGGAGAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 588  
 DB 61 GAACGAGAGGAGAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 QY 589 ATGAGATAGTCCAGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 648  
 DB 121 ATGAGATAGTCCAGAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
 QY 649 ACCTGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 708  
 DB 181 ACCTGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
 QY 709 GGCCTACACCGGGGCTCTCTTCCAGACACACAGAGAGAGTGGGCGGCTCTTGTG 768  
 DB 241 GGCCTACACCGGGGCTCTCTTCCAGACACACAGAGAGAGTGGGCGGCTCTTGTG 300  
 QY 769 GAGGGAGACCGGCTCCACACACACACACACACACACACACACACACACAC 828  
 DB 301 GAGGGAGACCGGCTCCACACACACACACACACACACACACACACACACAC 861  
 QY 829 AAAGAGAGGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861  
 DB 361 AAAGAGAGGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393

RESULT 15  
 LOCUS BM714472 633 bp mRNA linear EST 28-FEB-2002  
 DEFINITION UI-E-EJ0-ahs-f-20-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone  
 ACCESSION BM714472  
 VERSION UI-E-EJ0-ahs-f-20-0-UI 5', mRNA sequence.  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 633)  
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE  
COMMENT

97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 593-614, >AT-rich#Low\_complexity (matched compliment)  
Seq primer: M13 Reverse.  
Location/Qualifiers

FEATURES  
source

1..633  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-abs-f-20-0-UI"  
/clone\_1db="UI-E-EJ0"  
/tissue\_type="Fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina Foveal and Macular, RPE and  
choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10b (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
UI-E-EJ0 is a subtracted cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT7T3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dr)18 tail. The  
sequence tags for this library are: fetal eyes, AGAATCAAGA  
; lens, CGATTAGCGA; eye anterior segment, AATGCCGCAT;  
optic nerve, CCATTAGGTG; retina, CCGCG; Retina Foveal and  
Macular, GTCC; RPE and Choroid, ACCCA. This library was  
created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."

BASE COUNT  
ORIGIN

150 a 206 c 182 g 93 t 2 others

Query Match 45.2%; Score 389; DB 14; Length 633;

Best local Similarity 100.0%; Pred. No. 4.4e-160;

Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 473 CCCGACCTTACGAGGCGCCCTGACCCCTTCAGCTTGGGACCCCGAGCAGCAGCTGGAAAC 532  
|||  
Db 14 CCCGACCTTACGAGGCGCCCTGACCCCTTCAGCTTGGGACCCCGAGCAGCAGCTGGAAAC 73  
0Y 533 TGAACCGGGAGTGTGCGCGGACCCCAACGAACATCTTTCGACAGTGCCTGATGG 592  
|||  
Db 74 TGAACCGGGAGTGTGCGCGGACCCCAACGAACATCTTTCGACAGTGCCTGATGG 133  
0Y 593 ATAGTCCAGAGGTGGGCGGCGCCCTGCGCCCGCAGAGTAATCGGGCATCAGGCCACGT 652  
|||  
Db 134 ATAGTCCAGAGGTGGGCGGCGCCCTGCGCCCGCAGAGTAATCGGGCATCAGGCCACGT 193  
0Y 653 GCTACGGCAGCGGCGGCGCATGTGAGGGGCGCCCGCCACCTACAGCGAGTCAATCGGCC 712  
|||  
Db 194 GCTACGGCAGCGGCGGCGCATGTGAGGGGCGCCCGCCACCTACAGCGAGTCAATCGGCC 253  
0Y 713 ACTACCGGGGGTCTCTTTCACACACAGAGAGAGTGGCGCCCTCTTGTGTGAGG 772  
|||  
Db 254 ACTACCGGGGGTCTCTTTCACACACAGAGAGAGTGGCGCCCTCTTGTGTGAGG 313  
0Y 773 GGACCGGGGTCCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAG 832

Db 314 |||||  
GGACCGGGGTCCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAG 373  
0Y 833 AGAAGATTAACAGAAAGAGACACCTCTC 861  
|||  
Db 374 AGAAGATTAACAGAAAGAGACACCTCTC 402

Search completed: March 18, 2003, 08:14:53  
Job time : 1303 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 12:10:39 ; Search time 2671.78 Seconds

(without alignments)  
9378.574 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861  
Sequence: 1 atgcacgcctgtgtggggf.....aacagaaaggacacccctc 861Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

```
GenBml:.*
1: gb_da:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*
29: em_vi:.*
30: em_htg_hum:.*
31: em_htg_inv:.*
32: em_htg_other:.*
33: em_htg_mus:.*
34: em_htg_pln:.*
35: em_htg_rtd:.*
36: em_htg_mam:.*
37: em_htg_vrt:.*
38: em_sy:.*
39: em_htgo_hum:.*
40: em_htgo_mus:.*
41: em_htgo_other:.*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	861	100.0	861	6 AX392419	AX392419 Sequence
2	861	100.0	1321	6 AX392417	AX392417 Sequence
3	861	100.0	4839	6 AF305616	AF305616 Homo sapi
4	839.4	97.5	1061	9 BC015918	BC015918 Homo sapi
5	754.2	87.6	1141	9 AF224278	AF224278 Homo sapi
6	752	87.3	1818	9 AV128643	AV128643 Homo sapi
7	612.4	71.1	878	6 AX392428	AX392428 Sequence
8	546.4	63.5	61505	9 AF305426	AF305426 Homo sapi
9	546.4	63.5	130435	9 HS718J7	AL035541 Human DNA
10	445.6	51.8	693	6 AX392430	AX392430 Sequence
11	421	48.9	651	10 AF220208	AF220208 Mus muscu
12	417.4	48.5	812	6 AX011709	AX011709 Sequence
13	401.4	46.6	408	6 AX071267	AX071267 Sequence
14	358	41.6	156075	2 AL837509	AL837509 Mus muscu
15	352.2	40.9	8494	9 AF009424	AF009424 Homo sapi
16	350	40.7	2170	9 AK055028	AK055028 Homo sapi
17	349	40.5	8093	6 AX392432	AX392432 Sequence
18	349	40.5	8093	9 AF009426	AF009426 Homo sapi
19	341.8	39.7	475	6 AX392431	AX392431 Sequence
20	284.2	33.0	8440	9 AF009425	AF009425 Homo sapi
21	281	32.6	8039	9 AF009427	AF009427 Homo sapi
22	250.2	29.1	172692	2 AP001013	AP001013 Homo sapi
23	250.2	29.1	173709	9 AP001010	AP001010 Homo sapi
24	250.2	29.1	181083	2 AP001268	AP001268 Homo sapi
25	240.8	28.0	766	10 BC022716	BC022716 Mus muscu
26	232.6	27.0	66972	2 AC117364	AC117364 Rattus no
27	231	26.8	183619	2 AC111069	AC111069 Mus muscu
28	228.2	26.5	150542	2 AC097603	AC097603 Rattus no
29	215.8	25.1	167758	2 AC110189	AC110189 Mus muscu
30	157.4	18.3	150224	9 HSJ1059L7	AL121913 Human DNA
31	152.2	17.7	167758	2 AC110189	AC110189 Mus muscu
32	120	13.9	159824	2 AC111878	AC111878 Rattus no
33	62	7.2	125020	9 AF429315	AF429315 Homo sapi
34	61.8	7.2	125020	9 AF429315	AF429315 Homo sapi
35	60.6	7.0	136452	2 AC111699	AC111699 Rattus no
36	59.8	6.9	88203	5 AC097628	AC097628 Takifugu
37	59.4	6.9	152983	2 AC094473	AC094473 Rattus no
38	58.8	6.8	219952	2 AC084804	AC084804 Mus muscu
39	58	6.7	141892	2 AC023197	AC023197 Mus muscu
40	57.8	6.7	26537	2 AC087228	AC087228 Mus muscu
41	56.6	6.6	236562	2 AL772338	AL772338 Mus muscu
42	56.2	6.5	45905	2 AC127121	AC127121 Rattus no
43	56.2	6.5	131364	2 AC128769	AC128769 Rattus no
44	55.8	6.5	300695	2 AC079431	AC079431 Mus muscu
45	55.6	6.5	166900	2 AC130935	AC130935 Rattus no

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AX392419	AX392419	Sequence 3 from Patent WO0216416.	AX392419	AX392419.1	GI:19700734	human.	Homo sapiens	Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and Turi,T.G.	I	Diagnosis and treatment of cardiovascular conditions
AX392419	AX392419	Sequence 3 from Patent WO0216416.	AX392419	AX392419.1	GI:19700734	human.	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and Turi,T.G.	I	Diagnosis and treatment of cardiovascular conditions





[illegible]

REFERENCE	1 (bases 1 to 1141)
AUTHORS	Xu, L. L., Shammugam, N., Segawa, T., Sesterhenn, I. A., McLeod, D. G., Moul, J. W., and Srivastava, S.
TITLE	A novel androgen-regulated gene, PMEPA1, located on chromosome 20q13 exhibits high level expression in prostate
JOURNAL	Genomics 66 (3), 257-263 (2000)
MEDLINE	20334621
PUBMED	10873380
REFERENCE	2 (bases 1 to 1141)
AUTHORS	Xu, L. L., Shammugam, N., Segawa, T., Sesterhenn, I. A., McLeod, D. G., Moul, J. W., and Srivastava, S.
JOURNAL	Direct Submission
REFERENCE	Submitted (12-JUN-2000) CDR, USUHS, 1530 East Jefferson Street, Rockville, MD 20852, USA
FEATURES	Location/Qualifiers
source	1. .1141
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="20"
	/map="20q13.31-13.33"
	/cell_line="LNCap"
	1. .1141
gene	/gene="PMEPA1"
	96. .854
CDS	/gene="PMEPA1"
	/note="Type 1b transmembrane protein; expression is induced in response to the synthetic androgen, R1881; expression is abundant in, and restricted to, prostatic glandular epithelial cells; similar to the predicted protein encoded by sequence deposited at GenBank Accession Number NP_004329"
	/codon_start=1
	/product="PMEPA1 protein"
	/protein_id="AAF66322.1"
	/db_xref="GI:9255809"
	/translation="MADEEFYQIIIIIVVMMVVVITCLLSHKLSARSPISHSOGRRRDLALSSECLMPSESTVSGNIPPEOVYAPRPDRLAVPFAERHRPOTPYRPLDGLPPTISDGEPPPYQGCTQLRLPBDLELNRESVAPRRTLPDSDLMDBARLGPGPCPPSSNSGISATCYGSGGRRMGPPPTSEVIGHYPGSSFFHQOQSGPPSLTGSTRHHHHHPLIESAAIWSKEMOKOKHPL"
BASE COUNT	271 a 350 c 336 g 184 t
ORIGIN	
Query Match	87.6%; Score 754.2; DB 9; Length 1141;
Best Local Similarity	98.3%; Pred. No. 9.3e-120;
Matches 762; Conservative	0; Mismatches 13; Indels 0; Gaps 0;
87	TTTGTTCAGAGCATGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCTGCT 146
77	TCTCTGGGAAACGAGCAATGGCCGAGCTGGAGTTTGTTCAGATCATCATCTGCT 136
147	GGTGATGATGGGATGGTGGTGGTATGATACCGTGGCTGAGCCACTCAAGCTGTGC 206
137	GGTGATGATGGTGGTGGTGGTGGTATGATACCGTGGCTGAGCCACTCAAGCTGTGC 196
207	ACGGCTCTTCATCAGCCGCGACAGCCAGGCGGAGAGAGAAGATGCCCTGTCTCAGA 266
197	ACGGCTCTTCATCAGCCGCGACAGCCAGGCGGAGAGAGAAGATGCCCTGTCTCAGA 256
267	AGGATGCTGTGGCCCTGGGAGAGCAGATGTTCAGGCAAGGAATCCAGAGCCGAGGT 326
257	AGGATGCTGTGGCCCTGGGAGAGCAGATGTTCAGGCAAGGAATCCAGAGCCGAGGT 316
327	CTACGCCCCGCTCGGCGCCAGCAGCCGCTGGCGTGGCGCCCTTCGGCCAGCGGAGG 386
317	CTACGCCCCGCTCGGCGCCAGCAGCCGCTGGCGTGGCGCCCTTCGGCCAGCGGAGG 376
387	CTTCCACCGCTTCAGCCACCTATTCGATCTGAGCAGAGATGACCTGGCCACCCAC 446
377	CTTCCACCGCTTCAGCCACCTATTCGATCTGAGCAGAGATGACCTGGCCACCCAC 436
447	CATTCGCTGTGAGACGGGAGGAGCCCGCCACCTTACCAAGGGCCCTGACCTCTCAGCT 506

D	b	437	CATCHCGCTGTCACAGGGGGAAGAGGCCCCCAOCCCTACAGAGGCCCTCACCCTTCCACT	496
O	y	507	TTCGGAGCCCCGAGCAGCAGCTTGAACTGAAACCGGAGATCGGTGCAGCACCCCMAAAG	566
D	b	497	TTCGGAGCCCCGAGCAGCAGCTTGAACTGAAACCGGAGATCGGTGCAGCACCCCMAAAG	556
O	y	567	AACCATCTTTCCAGCACTGACCTGATGGAATAAGTCCAGAGCTGGGGGGCCCCCTGCCCCCAG	626
D	b	557	AACCATCTTTCCAGCACTGACCTGATGGAATAAGTCCAGAGCTGGGGGGCCCCCTGCCCCCAG	616
O	y	627	CAGTAACCTGGGCACTACAGCGCCACGTGCTACAGCGCAGCGGCGGCCTATGGAAGGGCGCC	686
D	b	617	CAGTAACCTGGGCACTACAGCGCCACGTGCTACAGCGCAGCGGCGGCCTATGGAAGGGCGCC	676
O	y	687	GCCCACTACAGCGAGTGCATCGGGCACAACCCGGGGTCTCTTCCAGACACAGCAGAG	746
D	b	677	GCCCACTACAGCGAGTGCATCGGGCACAACCCGGGGTCTCTTCCAGACACAGCAGAG	736
O	y	747	CAGTGGCGCCGCTCTTCTTGTGTGAGGGGACCGGCTCCACACACACATCGGCGCCCT	806
D	b	737	CAGTGGCGCCGCTCTTCTTGTGTGAGGGGACCGGCTCCACACACACATCGGCGCCCT	796
O	y	807	AGAGAGCGCAGCCATCTGGAGCAAGAAGATTAACAAGAAAGGACACCCCTTC	861
D	b	797	AGAGAGCGCAGCCATCTGGAGCAAGAAGATTAACAAGAAAGGACACCCCTTC	851
<b>RESULT 6</b>				
LOCUS	AY128643	1818 bp	mRNA	linear PRI 05-AUG-2002
DEFINITION	Homo sapiens PMEPAL variant A protein mRNA, complete cds.			
ACCESSION	AY128643			
VERSION	AY128643.1	GI:22121998		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1818)			
AUTHORS	Brunschwig,E.B., Wilson,K., Mack,D.D., Dawson,D., Lawrence,E.,			
	Willson,J.K.V., Lu,S., Nosrati,A., Swinler,S., Beard,L.,			
	Luttrebaugh,J.D., Willis,J., Plazer,P. and Markowitz,S.			
TITLE	Direct Submission			
JOURNAL	Submitted (03-JUL-2002) Department of Medicine, Case Western			
	Reserve University/Howard Hughes Medical Institute, 11001 Cedar			
	Ave., Cleveland, OH 44106, USA			
FEATURES	Location/Qualifiers			
SOURCE	1..1818			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="20"			
	/map="20q13.31-13.33"			
	145..858			
	/codon_start=1			
	/product="PMEPAL variant A protein"			
	/protein_id="AA089277.1"			
	/db_xref="GI:22121999"			
	/translation="MMYVVVITCLLSHYKLARSFISRSQGRREDLSSSECLMP			
	SESIVSGNGIDPEOVYAAPRPETDLIAPPKQRRPRFPQTYLYOHETDLPRTSL			
	SDGEPPPYOGPCFLQLRDPDOOLELNRESYRAPNRTIFDSLDMSARLAGCPRESS			
	NSGISATCYGSYGMEGPPTYSVIIGHYPSSFOHOOSQPSLLLEGTRLHTHIAP			
	LESVAIWSKEKDOKRGHP."			
BASE COUNT	459 a 467 c 494 g 398 t			
ORIGIN				
Query Match	87.3%; Score 752; DB 9; Length 1818;			
Best Local Similarity	100.0%; Pred. No. 2.le-119;			
Matches 752; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
Oy	110 CGGACCTGGAGTTGTTCAGATCATCATTCGATCGTGTGTGATGATGATGGTGGTGG	169		
D	b 104 CGGACCTGGAGTTGTTCAGATCATCATTCGATCGTGTGTGATGATGATGGTGGTGG	163		



```

LOCUS       AF305426                61505 bp    DNA        linear    PRI 12-OCT-2001
DEFINITION   Homo sapiens solid tumor-associated 1 protein (STAG1/PMEP1) gene,
ACCESSION   AF305426
VERSION     AF305426.1    GI:15824468
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 61505)
AUTHORS     Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
TITLE       Identification and characterization of a novel gene, STAG1,
            up-regulated in renal cell carcinoma and other solid tumours
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 61505)
AUTHORS     Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
TITLE       Direct Submission
JOURNAL     Submitted (13-SEP-2000) Centre for Molecular Biotechnology,
            Queensland University of Technology, 2 George St, Brisbane, QLD
            4001, Australia

FEATURES
             source
             location/Qualifiers
                1..61505
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="20"
                /map="20q13.2-q13.33"
                1..61505
                /gene="STAG1/PMEP1"
                join(1..429,50206..50360,56817..56870,57305..61505)
                /gene="STAG1/PMEP1"
                /product="solid tumor-associated 1 protein"
                join(321..429,50206..50360,56817..56870,57305..57850)
                /gene="STAG1/PMEP1"
                /codon_start=1
                /product="solid tumor-associated 1 protein"
                /protein_id="AL09357.1"
                /db_xref="GI:15824469"
                /translation="MHRIMGVNSTAAAGOPVSGTCNCKRSLRSMETLEFVOI
                ITIVVMVWVYITCLISHYKISARSPISRHQGRREDALISSECLMPSESTVGN
                GIDPEQVIAVPRDRLAVPEAFQRRERHFOPTYLQHEIDLPTISLSDEEPP
                YGGPCLQRLDPEQQLNRESVRAPEKNTIFPDSLDMSARLGPCPPSNGISATC
                YGSGGMEGPPPYSEVIGHYPPSSFOHQSSGPPSLRSTRLHHTHIALPESAAIMS
                KERDKOKGHP."
BASE COUNT   13329 a 16074 c 17430 g 14672 t
ORIGIN
Query Match      63.5%; Score 546.4; DB 9; Length 61505;
Best Local Similarity 98.9%; Pred. No. 2,4e-84;
Matches 550; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY  306  CGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCCACGACGCGCTGCGCGGCC 365
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57292 CTGCTTCTCCAGCGGAGGTCTACGCCCGCTCGGCCACGACGCGCTGCGCGGCC 57351

QY  366  GCCCTTGGCCAGGAGGAGGCTTCACACGCTTCACGCCACCATTCGCTGACGAGA 425
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57352 GCCCTTGGCCAGGAGGAGGCTTCACACGCTTCACGCCACCATTCGCTGACGAGA 57411

QY  426  CGAGATGACCTGCACCCACACATCTGCTGTACAGAGGGGAGAGAGGCCACCTTACCA 485
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57412 CGAGATGACCTGCACCCACACATCTGCTGTACAGAGGGGAGAGAGGCCACCTTACCA 57471

QY  486  GGGGCCCTTCACCTTCACGCTTGGGAGCCCGGACAGCAGCTGGAAGTGAACGGGAGTC 545
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57472 GGGGCCCTTCACCTTCACGCTTGGGAGCCCGGACAGCAGCTGGAAGTGAACGGGAGTC 57531

QY  546  GGTCGGCGACCCCAAGAACATCTTCGACAGTGACCTGATGATGATGTCACAGCT 605
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57532 GGTCGGCGACCCCAAGAACATCTTCGACAGTGACCTGATGATGATGTCACAGCT 57591

QY  606  GGGCGGCCCTTCGCCCCCGGACAGCACTGAGGAGCCACGTCGTAACGGCAGCGG 665
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

Db  57592 GGGCGGCCCTTCGCCCCCGGACAGCACTGAGTACGGGATCAGGCCACGTCACGACGCG 57651
QY  666  CGGGCGCATGAGAGGGGGCGGCCCGCCACTTACAGGAGGTATGAGGCACACTACCCGGGCTC 725
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57652 CGGGCGCATGAGAGGGGGCGGCCCGCCACTTACAGGAGGTATGAGGCACACTACCCGGGCTC 57711

QY  726  CTCCTTCACAGCAGAGAGAGAGTGGGCGCCCTCTCTGTGAGGGAGCCCGGCTCCA 785
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57712 CTCCTTCACAGCAGAGAGAGAGTGGGCGCCCTCTCTGTGAGGGAGCCCGGCTCCA 57771

QY  786  CCACACACATGCGCGCCCTAGAGAGCGGACCATCTGAGCAAAAGAGATTAACA 845
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57772 CCACACACATGCGCGCCCTAGAGAGCGGACCATCTGAGCAAAAGAGATTAACA 57831

QY  846  GAAAGACACCCCTCTC 861
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db  57832 GAAAGACACCCCTCTC 57847

RESULT  9
LOCUS     HS718U7/c
DEFINITION
Human DNA sequence from clone RP4-718U7 on chromosome
20q13.31-13.33 Contains the PKI gene for soluble
phosphoenolpyruvate carboxylase 1, part of a novel gene similar
to mouse DLM-1 (tumor stroma and activated macrophage protein),
the 3' end of the TMPEP1 gene encoding an androgen induced lb
transmembrane protein (PMEP1), two putative novel genes, a Cpg
island, ESTs, STS and GSSs, complete sequence.
AL035541
AL035541.15 GI:11546043
HTG: Cpg island: DLM-1: macrophage protein: PKI:
phosphoenolpyruvate carboxylase; PMEP1; TMPEP1; transmembrane
protein.
human.
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 130435)
Sehra,H.
Direct Submission
Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 5, 2000 this sequence version replaced gi:10198628.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
This sequence is the entire insert of clone RP4-718U7 The true left
end of clone RP5-1007B6 is at 71437 in this sequence. The true
right end of clone RP4-579F20 is at 43945 in this sequence.
RP4-718U7 is from the library RPCR-4 constructed by the group of
Pleier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: PCYPAC2
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the

```



```
FEATURES
source      assembly was confirmed by restriction digest.
            Location/Qualifiers
            1. .130435
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="20"
               /map="q13.31-13.33"
               /clone="RP4-718J7"
               /clone_11b="RPC1-4"
            319. .478
               /note="5 copies 32 mer 67% conserved"
            repeat_region
            370. .463
               /note="2 copies 47 mer 87% conserved"
            repeat_region
            427. .626
               /note="4 copies 50 mer 71% conserved"
            repeat_region
            707. .812
               /note="53 copies 2 mer cc 61% conserved"
            repeat_region
            967. .2752
               /note="893 copies 2 mer 99 54% conserved"
            repeat_region
            982. .2730
               /note="33 copies 53 mer 54% conserved"
            repeat_region
            1177. .1820
               /note="4 copies 161 mer 64% conserved"
            repeat_region
            1204. .1371
               /note="3 copies 56 mer 75% conserved"
            repeat_region
            1358. .2757
               /note="28 copies 50 mer 54% conserved"
            repeat_region
            1675. .2718
               /note="18 copies 58 mer 55% conserved"
            repeat_region
            1843. .2346
               /note="9 copies 56 mer 64% conserved"
            repeat_region
            1852. .2079
               /note="4 copies 57 mer 86% conserved"
            repeat_region
            2028. .2632
               /note="11 copies 55 mer 60% conserved"
            repeat_region
            2087. .2257
               /note="3 copies 57 mer 83% conserved"
            repeat_region
            2228. .2497
               /note="5 copies 54 mer 75% conserved"
            repeat_region
            2583. .2716
               /note="2 copies 67 mer 82% conserved"
            repeat_region
            3378. .3426
               /note="LIMB4 repeat: matches 6088. .6136 of consensus"
            repeat_region
            3799. .3896
               /note="Charliel repeat: matches 681. .781 of consensus"
            repeat_region
            5331. .5793
               /note="MUR1B repeat: matches 14. .466 of consensus"
            repeat_region
            5797. .5988
               /note="6 copies 32 mer 86% conserved"
            repeat_region
            5855. .5978
               /note="31 copies 4 mer gcac 61% conserved"
            repeat_region
            6474. .6591
               /note="MIR repeat: matches 91. .218 of consensus"
            repeat_region
            6592. .6723
               /note="FLAM_C repeat: matches 1. .132 of consensus"
            repeat_region
            6952. .7021
               /note="L2 repeat: matches 2637. .2705 of consensus"
            repeat_region
            7358. .7671
               /note="AlusX repeat: matches 1. .312 of consensus"
            repeat_region
            8521. .8554
               /note="Alu repeat: matches 1. .34 of consensus"
            repeat_region
            8863. .9217
               /note="Alu repeat: matches 105. .505 of consensus"
            repeat_region
            9228. .9746
               /note="L2 repeat: matches 1459. .2028 of consensus"
            repeat_region
            9880. .10089
               /note="L2 repeat: matches 2292. .2511 of consensus"
            repeat_region
            10094. .10206
               /note="AluX repeat: matches 188. .300 of consensus"
            repeat_region
            10263. .10573
               /note="AluX repeat: matches 1. .308 of consensus"
            repeat_region
            10658. .10753
               /note="L2 repeat: matches 2647. .2749 of consensus"
            repeat_region
            11816. .12380
               /note="LTR19B repeat: matches 1. .580 of consensus"
            repeat_region
            12519. .12813
               /note="LIMD2 repeat: matches 6032. .6331 of consensus"
            repeat_region
            13559. .13726
               /note="3 copies 56 mer 73% conserved"
            repeat_region
            13585. .13716
               /note="22 copies 6 mer ctctct 67% conserved"
            repeat_region
            13586. .13717
               /note="66 copies 2 mer tc 68% conserved"
            repeat_region
            13588. .13715
               /note="4 copies 32 mer 78% conserved"
            repeat_region
            13589. .13724
               /note="34 copies 4 mer cttt 77% conserved"
            repeat_region
            14313. .14427
               /note="match: GSS: Em:AO592603"
            repeat_region
            14313. .14427
               /note="LTR16C repeat: matches 257. .386 of consensus"
            repeat_region
            15015. .15110
               /note="3 copies 32 mer 79% conserved"
            repeat_region
            15021. .15110
               /note="45 copies 2 mer ac 80% conserved"
            repeat_region
            15026. .15109
               /note="7 copies 12 mer 82% conserved"
            repeat_region
            15028. .15111
               /note="14 copies 6 mer cacaca 82% conserved"
            repeat_region
            15029. .15108
               /note="20 copies 4 mer acac 83% conserved"
            repeat_region
            15273. .15399
               /note="L2 repeat: matches 2159. .2285 of consensus"
            repeat_region
            15420. .15721
               /note="AlusX repeat: matches 1. .303 of consensus"
            repeat_region
            16333. .16396
               /note="2 copies 32 mer 98% conserved"
            repeat_region
            16575. .16690
               /note="L2 repeat: matches 2596. .2711 of consensus"
            repeat_region
            18049. .18169
               /note="MIR repeat: matches 86. .211 of consensus"
            repeat_region
            18312. .18438
               /note="MIR repeat: matches 138. .250 of consensus"
            repeat_region
            18585. .18776
               /note="MER20 repeat: matches 7. .217 of consensus"
            repeat_region
            19211. .19290
               /note="MIR repeat: matches 72. .154 of consensus"
            repeat_region
            20413. .20463
               /note="MIR repeat: matches 90. .143 of consensus"
            repeat_region
            20413. .20463
               /note="match: GSS: Em:AO748384"
            repeat_region
            21672. .21672
               /note="match: GSS: Em:AO776209"
            repeat_region
            21663. .21663
               /note="match: GSS: Em:AO136459"
            repeat_region
            21672. .21672
               /note="match: GSS: Em:AO892114"
            repeat_region
            21667. .21775
               /note="MIR repeat: matches 73. .192 of consensus"
            repeat_region
            23577. .23646
               /note="2 copies 35 mer 100% conserved"
            repeat_region
            24257. .24591
               /note="L2 repeat: matches 2336. .2710 of consensus"
            repeat_region
            25580. .25611
               /note="16 copies 2 mer tc 87% conserved"
            repeat_region
            26333. .26643
               /note="AluX repeat: matches 1. .310 of consensus"
            repeat_region
            27603. .27684
               /note="L2 repeat: matches 2661. .2739 of consensus"
            repeat_region
            27823. .28041
               /note="MIR repeat: matches 38. .242 of consensus"
            repeat_region
            28129. .28259
               /note="FLAM_C repeat: matches 1. .127 of consensus"
            repeat_region
            31203. .31356
               /note="LIMB5 repeat: matches 6015. .6176 of consensus"
            repeat_region
            32224. .32522
```



/note="Alusx repeat: matches 1..298 of consensus"  
repeat\_region 33285..33481

Query Match 63.5% Score 546.4; DB 9; Length 130435;  
Best Local Similarity 98.9% Pred. No. 2.2e-84;  
Matches 550; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 306 CGGATCCCAAGGCGGCTAGTGGCCGCTGGCCGACGCGCCGCGCTGGCCGCTGC 365  
| | | | |  
Db 128454 CTGCTTCTCCAGCGGAGGTCTACGCCCCGCTGGCCGACGCGCTGGCCGCTGC 128395  
QY 366 GCCCTTCGCGGAGGAGCGCTTCACCGCTTCAGCCGCACTATCCGTAACCTGACAGA 425  
| | | | |  
Db 128394 GCCCTTCGCGGAGGAGCGCTTCACCGCTTCAGCCGCACTATCCGTAACCTGACAGA 128335  
QY 426 CGAGATGACCTGCCACCCACCATCTGCTGTACAGCGGGAGAGCCGCCACCTTACCA 485  
| | | | |  
Db 128334 CGAGATGACCTGCCACCCACCATCTGCTGTACAGCGGGAGAGCCGCCACCTTACCA 128275  
QY 486 GGGCCCTTCGACCGCTTCAGCTTCGGGACCCGAGCAGAGGTGGAACCGGGAGTGC 545  
| | | | |  
Db 128274 GGGCCCTTCGACCGCTTCAGCTTCGGGACCCGAGCAGAGGTGGAACCGGGAGTGC 128215  
QY 546 GGTGCGCGACCGCCCAACAGAACCATCTTCGACAGTACCTGATGATGTCAGGCT 605  
| | | | |  
Db 128214 GGTGCGCGACCGCCCAACAGAACCATCTTCGACAGTACCTGATGATGTCAGGCT 128155  
QY 606 GGGGCGCGCTTCGCGCCCGCCAGCATCTCGGCGATCAGCGCGCTCTACGGCAGCG 665  
| | | | |  
Db 128154 GGGGCGCGCTTCGCGCCCGCCAGCATCTCGGCGATCAGCGCGCTCTACGGCAGCG 128095  
QY 666 CGGGCGATGAGAGGGGCGCGCCGACCTACAGCAGAGGTGCACTGCGCGGCTGC 725  
| | | | |  
Db 128094 CGGGCGATGAGAGGGGCGCGCCGACCTACAGCAGAGGTGCACTGCGCGGCTGC 128035  
QY 726 CTCCTTCAGACGACGAGAGCATGAGCGCCGCTCTTGTGAGGGGACCGGCTTCA 785  
| | | | |  
Db 128034 CTCCTTCAGACGACGAGAGCATGAGCGCCGCTCTTGTGAGGGGACCGGCTTCA 127975  
QY 786 CCACACACACATCGCGCCCTTAGAGAGCGGCACTCTGGAGCAAGAGATTAACA 845  
| | | | |  
Db 127974 CCACACACACATCGCGCCCTTAGAGAGCGGCACTCTGGAGCAAGAGATTAACA 127915  
QY 846 GAAAGGACACCCCTTC 861  
| | | | |  
Db 127914 GAAAGGACACCCCTTC 127899

RESULT 10  
AX392430/c 693 bp DNA linear PAT 23-MAR-2002  
LOCUS AX392430  
DEFINITION Sequence 14 from Patent WO0216416.  
ACCESSION AX392430  
VERSION AX392430.1 GI:19700746  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Lee,R.T., Landschultz,K.T., Kennedy,S.P., Thompson,J.F. and  
Turti,T.G.  
TITLE Diagnosis and treatment of cardiovascular conditions  
JOURNAL THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)  
FEATURES  
source 1..693  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 101 a 205 c 237 g 149 t 1 others  
ORIGIN  
Query Match 51.8% Score 445.6; DB 6; Length 693;

Best Local Similarity 97.2% Pred. No. 7.5e-67;  
Matches 485; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 364 CGCCCTTCGCCAGCGGAGCGCTTCCAGCGCTTCAGCCGACCTATCCGTACCTGCAG 423  
| | | | |  
Db 693 CGCCCTTCGCCAGCGGAGCGCTTCCAGCGCTTCCAGCCGACCTATCCGTACCTGCAG 634  
QY 424 CAGGATGACCTGCCACCCACCATCTGCTGTACAGCGGGAGAGACCCGCCACCTTAC 483  
| | | | |  
Db 633 CAGGATGACCTGCCACCCACCATCTGCTGTACAGCGGGAGAGACCCGCCACCTTAC 577  
QY 484 CAGGCGCCCTTCACCGCTTCAGCTTCGGAGCCCGAGCAGAGCTGGAAGTGAACCGGAG 543  
| | | | |  
Db 576 CAGGCGCCCTTCACCGCTTCAGCTTCGGAGCCCGAGCAGAGCTGGAAGTGAACCGGAG 517  
QY 544 TCGGTGCGCGACCGCCCAACAGAACCATCTTCGACAGTGAACCTGATGATGTCAGG 603  
| | | | |  
Db 516 TCGGTGCGCGACCGCCCAACAGAACCATCTTCGACAGTGAACCTGATGATGTCAGG 457  
QY 604 CTGGGCGCGCCCTTCGCGCCCGCCAGCATCTGCGGATCAGCGGCGAGCGGAGC 663  
| | | | |  
Db 456 CTGGGCGCGCCCTTCGCGCCCGCCAGCATCTGCGGATCAGCGGCGAGCGGAGC 397  
QY 664 GGGGCGCATGAGAGGGGCGCGCCGACCTACAGCAGAGTTCATCGGCACTACCGGGG 723  
| | | | |  
Db 396 GGGGCGCATGAGAGGGGCGCGCCGACCTACAGCAGAGTTCATCGGCACTACCGGGG 337  
QY 724 TCTCTTCACAGCAGCAGCAGAGCATGAGGGCGCGCTCTCTTGTGGA-GGGACCGGCT 782  
| | | | |  
Db 336 TCTCTTCACAGCAGCAGCAGAGCATGAGGGCGCGCTCTCTTGTGGA-GGGACCGGCT 277  
QY 783 CCACACACACATCGCGCCCTTAGAGAGCGGACCATCTGGAGCAAGAGATTAACA 842  
| | | | |  
Db 276 CCACACACACATCGCGCCCTTAGAGAGCGGACCATCTGGAGCAAGAGATTAACA 217  
QY 843 ACAGAAAGGACACCCCTTC 861  
| | | | |  
Db 216 ACAGAAAGGACACCCCTTC 198

RESULT 11  
AF220208 651 bp mRNA linear ROD 05-MAR-2001  
LOCUS AF220208  
DEFINITION Mus musculus Nedd4 WW domain-binding protein 4 mRNA, partial cds.  
ACCESSION AF220208  
VERSION AF220208.1 GI:12004973  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 651)  
Jolliffe,C.N., Harvey,K.F., Haines,B.P., Parasivam,G. and Kumar,S.  
TITLE Identification of multiple proteins expressed in murine embryos as  
binding partners for the WW domains of the ubiquitin-protein ligase  
Nedd4

JOURNAL Biochem. J. 351 Pt 3, 557-565 (2000)  
MEDLINE 20498735  
PUBMED 11042109  
REFERENCE  
AUTHORS Jolliffe,C.N. and Kumar,S.  
TITLE Direct Submission  
JOURNAL Submitted (30-DEC-1999) Division of Haematology, Hanson Centre for  
Cancer Research, IMVS, Frome Rd, Adelaide, SA 5000, Australia  
FEATURES  
source 1..651  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_lib="Novagen Cat. No. 69640-3"  
/dev\_stage="16 day embryo"  
<1..609  
/note="N4MBP4"  
/codon\_start=1

CDS  
Query Match

/product="Nedd4\_WW domain-binding protein 4"  
 /protein\_id="AA644247.1"  
 /db\_xref="GI:12004974"  
 /translation="SOARRRDLISSEGLMPSESTVSGMPEPOYAPRPTDLAY  
 PPRFONRQPTPYLOHEITAPRTISLSDGEPYOGPCTLOLRDPEOLELRES  
 VRAPNRTIDSLIDSTMIGPCPPSSNSGSAISATVSSGGMPEGPPTYSIVIGHYP  
 GSSFOHOOSNGPSLSLEGTFLHSHIAPLENKEKEKOKRHP."
   
BASE COUNT 146 a 234 c 176 g 95 t  
 ORIGIN

Query Match 48.9%; Score 421; DB 10; Length 651;  
 Best Local Similarity 83.18; Pred. No. 1.2e-62;  
 Matches 526; Conservative 0; Mismatches 80; Indels 27; Gaps 3;

QY 229 AGCCAGGGGGGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 348  
 Db 1 AGCCAGGGGGGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 348  
 QY 289 AGCAGAGTGTGAGGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 348  
 Db 61 AGTATGCTGTGAGGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 348  
 QY 349 GAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 117  
 Db 118 GAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 117  
 QY 409 TATCCGTACTGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 168  
 Db 169 TATCCGTACTGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 168  
 QY 469 GAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528  
 Db 229 GAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528  
 QY 529 GAACTGAAACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288  
 Db 289 GAACTGAAACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288  
 QY 589 ATGATAGTGTGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 348  
 Db 349 ATGATAGTGTGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 348  
 QY 649 AGCTGCTGAGAGAGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 408  
 Db 409 AGCTGCTGAGAGAGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 408  
 QY 709 GAGCAGTACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468  
 Db 469 GAGCAGTACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468  
 QY 769 GAGGAGACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528  
 Db 529 GAGGAGACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528  
 QY 829 AAGAGAGAGATTAACAGAGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCC 861  
 Db 574 AAGAGAGAGATTAACAGAGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCC 861  
 RESULT 12 812 bp DNA Linear PAT 06-SEP-2000  
 AX011709  
 LOCUS  
 DEFINITION Sequence 107 from Patent W09955858.  
 ACCESSION AX011709  
 VERSION AX011709.1 GI:9998233  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 812)  
 AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and  
 Pilarsky,C.

TITLE Human nucleic acid sequences obtained from pancreas tumor tissue  
 JOURNAL  
 PATENT: WO 9955858-A 107 04-NOV-1999;  
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
 BERNARD (DE); ROSENTHAL ANDRE (DE); METZGER GUS FÜRER GENOMFORSCHUNG  
 (DE); PILARSKY CHRISTIAN (DE)  
 FEATURES  
 source location/Qualifiers  
 1..812  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 157 a 272 c 237 g 146 t  
 ORIGIN

Query Match 48.5%; Score 417.4; DB 6; Length 812;  
 Best Local Similarity 92.08; Pred. No. 5e-62;  
 Matches 461; Conservative 0; Mismatches 16; Indels 24; Gaps 1;

QY 47 GCGAGCCCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 106  
 Db 324 GCGAGCCCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 106  
 QY 107 TCACGAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 166  
 Db 384 TCACGAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 166  
 QY 167 TGGTGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425  
 Db 426 TGGTGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425  
 QY 227 ACAGCAGAGGGGGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 286  
 Db 480 ACAGCAGAGGGGGAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 286  
 QY 287 AGAGCAGAGTGTGAGGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 346  
 Db 540 AGAGCAGAGTGTGAGGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 346  
 QY 347 CGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406  
 Db 600 CGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406  
 QY 407 CCTATCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466  
 Db 660 CCTATCGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466  
 QY 467 AGAGACCGCCCACTTACAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 526  
 Db 720 AGAGACCGCCCACTTACAGAGAGAGATGCCCTGCTCTCAGAGAGAGAGCTGTGGCCCTGGCCAC 526  
 QY 527 TGGAGCTGAAACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547  
 Db 780 TGGAGCTGAAACCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547  
 RESULT 13 408 bp DNA Linear PAT 25-JAN-2001  
 AX071267  
 LOCUS  
 DEFINITION Sequence 1739 from Patent W00102568.  
 ACCESSION AX071267  
 VERSION AX071267.1 GI:12581618  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 408)  
 AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,  
 Kassam,A., Reinhard,C., Randoz,F., Kennedy,G.C., Pot,D.,  
 Lanson,G., Drmanac,R., Ckenjakov,R., Drmanac,S., Dickson,M.,  
 Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.  
 Human genes and gene expression products  
 PATENT: WO 0102568-A 1739 11-JAN-2001;  
 CITRON CORPORATION (US); HYSEQ, INC. (US)  
 location/Qualifiers

FEATURES



```

Db 134299 TTCCAGACGAGCAAAATGACGGGCGCATCCCTGCTAGAGGGGACCGGCTCCATCCAC 134240
QY 790 ACACACATCGGGCCCTAGAGAGCGGCGCATCTGGAGCAAGAGAGATAAACAGAAA 849
Db 134239 TCGACACTTGGCCCTGCTGGA-----GACAAAGAGAGAGAGAGAAACAGAAA 134195
QY 850 GGACACCCCTCTC 861
Db 134194 GGTACCCCTCTC 134183

RESULT 15
AF009424 8494 bp mRNA linear PRI 17-FEB-1998
LOCUS Homo sapiens clone 22 mRNA, alternative splice variant alpha-1,
DEFINITION complete cds.
ACCESSION AF009424
VERSION AF009424.1 GI:2271468
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 8494)
AUTHORS Yoshikawa,T., Sanders,A.R., Esterling,L.E., Overhauser,J.,
Garnes,J.A., Lennon,G., Grewal,R. and Detera-Wadleigh,S.D.
TITLE Isolation of chromosome 18-specific brain transcripts as positional
candidates for bipolar disorder
JOURNAL Am. J. Med. Genet. 74 (2), 140-149 (1997)
MEDLINE 97275951
REFERENCE 2 (bases 1 to 8494)
AUTHORS Yoshikawa,T., Sanders,A.R., Esterling,L.E. and Detera-Wadleigh,S.D.
TITLE Multiple transcriptional variants and RNA editing in C18orf1, a
novel gene with IDRA and transmembrane domains on 18p11.2
JOURNAL Genomics 47 (2), 246-257 (1998)
MEDLINE 98140124
REFERENCE 3 (bases 1 to 8494)
AUTHORS Yoshikawa,T. and Detera-Wadleigh,S.D.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-1997) Clinical Neurogenetics Branch, National
Institute of Mental Health, Bethesda, MD 20892, USA
FEATURES
source Location/Qualifiers
1..8494
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p11.2"
470..1390
/note="alternatively spliced; alpha-1 form; possible
membrane-spanning protein"
/product="clone 22"
/codon_start=1
/product="clone 22"
/db_xref="GI:2271469"
/translation="MPEAGFOATNAFTCKFTTSGKCLYLGLSVQNDQNDGNSDE
ENCLVTEHPPEGLFNSLEFAQIIIVVTVVWVYVCLNHYKSTRSFINRPNQ
SRREDGLPQEGCLPDSAPRLGASEIMHAPSRDFTAPSFQIORDRFSRPOPTP
YVOHEIDLPPTSLDGEPPPYOGPTCLORDPEQOMELNRESYRAPNRTIFDSOL
IDIAMYSGGPCPPSPNSGISAISCTSSNRMGEPPTYSYEVNGHHPGASFLHQRSMNH
RGRRIQFOONNAESTIVPIKGRKRPGLV"
BASE COUNT 2285 a 1900 c 1907 g 2400 t 2 others
ORIGIN
Query Match 40.9%; Score 352.2; DB 9; Length 8494;
Best Local Similarity 73.1%; Pred. No. 5.7e-51;
Matches 482; Conservative 0; Mismatches 166; Indels 9; Gaps 2;
QY 94 CAGACATGAGATCAGGAGCTGGAGTTGTTGATCATCATCATCATGATG 153
Db 635 CCGGCACTCTCACTCGAGAGCTGGATTCGCCCAATCATCATCATCATG 694

```

```

QY 154 ATGCTATGTTGTTGTTGATCACTGCTGCTGAGCCACTACAGCTGTCTGACAGTTC 213
Db 695 ACGGTATGTTGTTGTTGATCACTGCTGCTGAGCCACTACAGCTGTCTGACAGTTC 754
QY 214 TTTCATCAGCGGCGGACAGGAGGCGGAGAGAGAGATGCGCTGCTCAGAGAGATGC 273
Db 755 TTTCATCAGCGGCGGCGGACAGGAGGCGGAGAGAGAGATGCGCTGCGGAGAGTTC 814
QY 274 CTGTGCGCTTGGAGAGCAGCAGTGTCAAGGCAACGGAATCCAGAGCCGAGTTCAGGC 333
Db 815 CTGTGCGCTTGGAGAGCAGCAGTGTCAAGGCGGCGGAGTTCAGAGTTCAGAT 868
QY 334 CGGCTTCGCGGCGGCGGACAGGAGGCGGCTGCGGCTGCGGCTTGGCCAGGAGGCTTCAC 393
Db 869 GCGGCGGCTTCAGGAGGAGGAGGCTTCAGGAGGCGGCTTGTATCAAGAGGAGTTCAGG 928
QY 394 CGGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 453
Db 929 CGGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 988
QY 454 CTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 513
Db 989 CTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1048
QY 514 CCCGAGCAGCAGCTGGAAGTGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 573
Db 1049 CCTGAGCAGCAGCTGGAAGTGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1108
QY 574 TTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 630
Db 1109 TTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1168
QY 631 AACTGGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 690
Db 1169 AACTGGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228
QY 691 AACTGAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 749
Db 1229 AACTGAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1287

```

Search completed: March 17, 2003, 15:53:56  
Job time : 2748.78 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 10:38:48 ; Search time 234.388 Seconds  
(without alignments)  
8272.492 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861

Sequence: 1 atgacacgccttgatgggggt.....aacagaagaagacacctctc 861

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq.101002.\*

```
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	861	100.0	1321	24	ABK12137	Human cDNA encoding
2	755.2	87.7	969	21	AAA75151	cDNA encoding a hu
3	754.2	87.6	1140	24	ABK92120	Prostate cancer-as
4	753.6	87.5	969	21	AAA75163	cDNA clone encodin
5	753.6	87.5	969	21	AAA75164	cDNA clone encodin
6	753.6	87.5	969	21	AAA75165	cDNA clone encodin
7	752.6	87.4	1061	21	AAA47429	Sequence encoding
8	742.4	86.2	1066	22	AA157868	Human polynucleoti
9	712	82.7	1069	22	AA159654	Human polynucleoti

10	612.4	71.1	878	24	ABK12142	Mouse cDNA encoding
11	462.2	53.7	1713	21	AAA75152	cDNA encoding a mu
12	460.6	53.5	1713	21	AAA75166	cDNA clone encodin
13	460.6	53.5	1713	21	AAA75167	cDNA clone encodin
14	460.6	53.5	1713	21	AAA75168	cDNA clone encodin
15	445.6	51.8	693	24	ABK12143	Human MYR-1 homol
16	417.4	48.5	22	AA252964	Human prostate tum	
17	401.4	46.6	408	22	AA252964	Novel human polynu
18	352.2	40.9	8494	23	AA252964	DNA encoding novel
19	351.6	40.8	937	21	AA252964	Human secreted pro
20	350	40.7	920	24	ABK34251	Human cDNA for nov
21	349	40.5	8065	19	AAV38335	Manic-depressive i
22	349	40.5	8093	24	ABK12145	Human MYR-1 homol
23	341.8	39.7	475	24	ABK12144	Human MYR-1 homol
24	333.8	38.8	8103	23	AA27312	DNA encoding novel
25	320.4	37.2	1679	23	AA284503	DNA encoding novel
26	284.2	33.0	8440	23	AA27313	DNA encoding novel
27	284.2	33.0	8440	24	ABK83477	Human cDNA differe
28	281	32.6	8011	19	AAV38336	Manic-depressive i
29	252.6	29.3	8041	23	AA27310	DNA encoding novel
30	232	26.9	254	21	AAA41265	Human secreted exp
31	223.4	25.9	1428	23	AA292497	DNA encoding novel
32	223.4	25.9	2942	23	AA27313	DNA encoding novel
33	51.2	5.9	114955	20	AA253491	Human adenosine A1
34	51	5.9	51	22	AA289714	Human coding sequ
35	50.6	5.9	4920	23	ABL09461	Drosophila melanog
36	50.6	5.9	5139	23	ABL09373	Drosophila melanog
37	50.6	5.9	5556	23	ABL09460	Drosophila melanog
38	49	5.7	60	24	ABN40872	Human gene express
39	48.4	5.6	1337	20	AA217263	Human gene express
40	48.4	5.6	11304	22	AA103113	Human gene express
41	47.8	5.6	1359	20	AA217254	Human gene express
42	47.8	5.6	8438	15	AA273500	DNA encoding Pseud
43	47.4	5.5	1215	21	AA255808	S. lavendulae MltG
44	47.4	5.5	3015	20	AA232027	Human MYR1 relate
45	47.4	5.5	3015	22	AA290084	446970 cDNA clone.

#### ALIGNMENTS

```
RESULT 1
ID ABK12137 standard; cDNA; 1321 BP.
XX ABK12137;
AC ABK12137;
XX 05-JUN-2002 (first entry)
DT
DE Human cDNA encoding Mechanically Induced Vascular Receptor 1, MYR-1.
XX
KW Human; ss; gene; MYR-1; Mechanically Induced Vascular Receptor 1;
KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;
KW cardiac cell; anti-apoptotic; vascular endothelial cell;
KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;
KW heart failure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 413..1276
FT /tag= a
FT /product= "MYR-1"
FT /note= "This region is specifically claimed in claim 3"
```

WO200216416-A2.

28-FEB-2002.

21-AUG-2001; 2001WO-US26089.

22-AUG-2000; 2000US-227159P.

PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.  
 PA (PITZ) PFIZER INC.

Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;  
 WPI: 2002-280912/32.

DR P-PSDB; AAU78231.

PT Novel nucleic acid molecule encoding Mechanically Induced Vascular  
 Receptor-1 polypeptide, useful for treating cardiovascular diseases  
 Claim 2; Page 87-88; 105pp; English.

The invention relates to an isolated nucleic acid molecule encoding a  
 Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
 cardiac cell anti-apoptotic activity and fragments of it provided  
 they are not identical to Genbank sequences A176144.1, A1594390,  
 NM004338 and A0177461. Also included are expression vectors, host  
 cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
 MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity  
 with a candidate agent, where the molecule is a nucleic acid molecule  
 comprising MIVR-1, IEX-1, VDU-1, Brg-2 and Trs-11d or its  
 expression product, determining if the anti-apoptotic activity is  
 modulated and thereby identifying a modulator. The cardiac cell anti-  
 apoptotic molecules and nucleic acids of the invention are useful for  
 treating, diagnosing and monitoring progression of such diseases and  
 disorders as characterized by increased apoptotic cell-death of vascular  
 endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
 stroke, arteriosclerosis and heart failure. The present sequence  
 encodes human MIVR-1.

Sequence 1321 BP; 223 A; 493 C; 440 G; 165 T; 0 other;

Query Match 100.0%; Score 861; DB 24; Length 1321;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-167;  
 Matches 861; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGCTGATGAGGGGTCACACACCGCCGCCGCCGGGAGCCCAATGTC 60  
 DB 413 ATGACCGCTGATGAGGGGTCACACACCGCCGCCGCCGGGAGCCCAATGTC 60  
 QY 61 TCTGACAGTGCMACTGCAAAAGCTCTTGTTCACAGACATGAGATCAGAGCTGAG 120  
 DB 473 TCTGACAGTGCMACTGCAAAAGCTCTTGTTCACAGACATGAGATCAGAGCTGAG 120  
 QY 121 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 DB 533 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 QY 181 CTGCTGAGCACTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 592  
 DB 593 CTGCTGAGCACTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 592  
 QY 241 AGGAGAGAGAGTCCCTCTCTGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 652  
 DB 653 AGGAGAGAGAGTCCCTCTCTGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 652  
 QY 301 GGCAAGAGAGTCCAGAGCGCAGTCTAGCGCCGCTGCGCCACCGACCGCTGCGCC 712  
 DB 713 GGCAAGAGAGTCCAGAGCGCAGTCTAGCGCCGCTGCGCCACCGACCGCTGCGCC 712  
 QY 361 GTGCGCGCTTGGCGCCAGCGGAGAGCGCTTCCACCGCTTCCAGCGCCATCTGAGTCTG 420  
 DB 773 GTGCGCGCTTGGCGCCAGCGGAGAGCGCTTCCACCGCTTCCAGCGCCATCTGAGTCTG 420  
 QY 421 CAGCAGAGATGAGTCCAGCCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832  
 DB 833 CAGCAGAGATGAGTCCAGCCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832  
 QY 481 TACCAAGGCGCCCTGCAAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 892  
 DB 893 TACCAAGGCGCCCTGCAAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 892

QY 541 GAGTCGAGCGCGCCACCCCAACAGACCATCTTGCAGAGTACCTGATGATGATGTC 600  
 DB 953 GAGTCGAGCGCGCCACCCCAACAGACCATCTTGCAGAGTACCTGATGATGATGTC 600  
 QY 601 AGGCTGAGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1012  
 DB 1013 AGGCTGAGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1012  
 QY 661 AGCG 720  
 DB 1073 AGCG 720  
 QY 721 GAGTCGCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1132  
 DB 1133 GAGTCGCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1132  
 QY 781 CTCACACACACACACACATCGCCCGCTAGAGAGCGCAGCAGCAGCAGCAGCAGCAGCAG 1192  
 DB 1193 CTCACACACACACACACATCGCCCGCTAGAGAGCGCAGCAGCAGCAGCAGCAGCAGCAG 1192  
 QY 841 AACAGAAAGAGACACCCCTTC 861  
 DB 1253 AACAGAAAGAGACACCCCTTC 1273

# RESULT 2

AAA75151  
 ID AAA75151 standard; cDNA; 969 BP.

AC AAA75151;

DT 15-JAN-2001 (first entry)

DE cDNA encoding a human TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hemopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW cardiovascular disorder; spleen associated disease; renal disorder;  
 KW brain herniation; latrogenic disease; ischemic heart disease; hydrocephalus;  
 KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 6..764

FT /tag- a

FT sig\_peptide /product- "TANGO 261"

FT mat\_peptide /tag- b

FT /tag- c

PN W0200052022-A1.

PD 08-SEP-2000.

PF 01-MAR-2000; 2000MO-US05226.

PR 01-MAR-1999; 990S-0122458.

PA (MILL-) MILLENNIUM PHARM INC.

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

DR WPI: 2000-579269/54.

DR P-PSDB; AAB18449.

XX Novel human and murine secreted proteins designated TANGO 216, 261,

PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PR e.g. for treating cancer -  
 XX  
 PS Claim 2, Fig 5, 175pp; English

XX

CC The present sequence encodes a human TANGO 261 polypeptide. The  
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and  
CC TANGO 267. The TANGO polypeptides can be used to modulate cellular  
CC proliferation, modulate cellular differentiation and/or modulate  
CC cellular adhesion. The proteins can be used to treat any von Willebrand  
CC factor-associated disorder, regulate extracellular matrix structuring,  
CC cellular adhesion, and cell trafficking and/or migration, modulate  
CC cellular interactions, modulate cell adhesion in proliferative  
CC disorders, such as cancer, modulate the proliferation, differentiation,  
CC and/or function of cells that appear in the bone marrow, and leukocytes,  
CC treat bone marrow, blood and hematopoietic associated diseases and  
CC disorders, alelectasis, pulmonary congestion or oedema, emphysema,  
CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
CC disorders, spleen associated diseases, modulate renal disorders, treat  
CC cardiovascular disorders such as ischemic heart disease, modulate the  
CC proliferation, differentiation, and/or function of bone and cartilage  
CC cells and to treat bone and/or cartilage associated diseases or  
CC disorder. They may also be used to treat disorders associated with the  
CC ovaries, cerebral oedema, hydrocephalus, brain herniations, latrogenic  
CC disease, inflammations, bacterial and viral meningitis, Alzheimer's  
CC disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis  
CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
CC disorders.

Sequence 969 BP; 211 A; 316 C; 294 G; 148 T; 0 other;

Query Match	87.7%	Score 755.2	DB 21	Length 969
Best Local Similarity	99.6%	Pred. NO. 1.1e-145		
Matches 757; Conservative	0	Mismatches 3		Indels 0
				Gaps 0

OY	102	GGAGATCAACGGAGCGTGGAGATTGTTACATATCATATCATATCGTGGTGGATGATGGAT	163
Db	2	GGAGATGGCGGAGCTGGAGATTGTTCAATATATCATCTATCATCTGGTGGTGGATGGAT	61
OY	162	GGTGGTGGTGCATCATGTCCTCTGTAGGCACTACAAAGTGTCTGCACAGTCTCTTCACAG	223
Db	62	GGTGGTGGTGCATCATGTCCTCTGTAGGCACTACAAAGTGTCTGCACAGTCTCTTCACAG	122
OY	222	CCGGCAACAGCCAGGGGGGGGAGAGAGAAGATGCCCTGTCTCAAGAAAGATGCCCTGTGCC	283
Db	122	CCGGCAACAGCCAGGGGGGGGAGAGAGAAGATGCCCTGTCTCAAGAAAGATGCCCTGTGCC	183
OY	282	CTCGAGAGAGCAACAGTGTAGGCAACGAAATCCAGAGCCGAGGCTACAGCCGCCGCCCTCG	343
Db	182	CTCGAGAGAGCAACAGTGTAGGCAACGAAATCCAGAGCCGAGGCTACAGCCGCCGCCCTCG	243
OY	342	GCCCAACGACCGCCCTGGGCGCTGCCCTTCGCCAGCGGGAGCGCTTCACACCGCTTCA	403
Db	242	GCCCAACGACCGCCCTGGGCGCTGCCCTTCGCCAGCGGGAGCGCTTCACACCGCTTCA	303
OY	402	GCCCAACGATCCGTATCCGTGCAGACAGCAATGACCTGCACCAACATCTGGCTGTAGA	463
Db	302	GCCCAACGATCCGTATCCGTGCAGACAGCAATGACCTGCACCAACATCTGGCTGTAGA	363
OY	462	CGGGGAGGAGCCCCCACCCTTACACAGGGGCCCTGTGACCTTCAGCTTGGGAGCCCGAGCA	523
Db	362	CGGGGAGGAGCCCCCACCCTTACACAGGGGCCCTGTGACCTTCAGCTTGGGAGCCCGAGCA	423
OY	522	GCAGCTGGAAGTGAACCGGGAGTGGGTGCGCGCACCCCAACAGAAACATCTTGCACAG	583
Db	422	GCAGCTGGAAGTGAACCGGGAGTGGGTGCGCGCACCCCAACAGAAACATCTTGCACAG	483
OY	582	TGACCTGATGATAGTATGCGCAGAGCTGGGGGCCGACCCCAACAGAAACATCTTGCACAG	643
Db	482	TGACCTGATGATAGTATGCGCAGAGCTGGGGGCCGACCCCAACAGAAACATCTTGCACAG	543
OY	642	CAGCGCAACGTGCTACGGCAGGGGCGGAGTGGAGGGGCGCGCCACCTACAGCA	703

Dd	542	CAGCGCCACGCTGTCTAAGCGCAACGCCGGCGGCAGCATGAAGGGCCCCCGCCCACTTACAGCCA	601
Oy	702	GGTATCTCGGCACACTACCOCGGGGGTCTCTTCAGCACCAGCAGAGAGTGGGCCGCCCTC	761
Dd	602	GGTATCTCGGCACACTACCOCGGGGGTCTCTCTTCAGCACCAGCAGAGAGTAGTGGCGGCCCTC	661
Oy	762	CTTGGCTGGAGGGGAGCCCGGCTCCACACACACACATCGCGCCCCTTAGAGAGCGCACCCAT	821
Dd	662	CTTGGCTGGAGGGGAGCCCGGCTCCACACACACACATCGCGCCCCTTAGAGAGCGCACCCAT	721
Oy	822	CTGAGAGCAAAGAAGATTAACAAGAAAGACACACCCCTC	861
Dd	722	CTGGAGCAAAGAAGATTAACAAGAAAGACACACCCCTC	761
<b>RESULT 3</b>			
ID	ABK92120	standard; DNA; 1140 BP.	
XX	ABK92120;		
AC	15-AUG-2002	(first entry)	
Df			
XX		Prostate cancer-associated DNA sequence #6.	
DE			
XX		Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;	
KW		gene therapy; gene; ds.	
KM			
XX	Mammalia.		
OS			
XX	WO200230268-A2.		
PN			
PD	18-APR-2002.		
XX			
PF	12-OCT-2001; 2001MO-US32045.		
XX			
PR	13-OCT-2000; 2000US-0687576.		
PR	08-DEC-2000; 2000US-0733288.		
PR	08-DEC-2000; 2000US-0733742.		
PR	24-JAN-2001; 2001US-263957P.		
PR	16-MAR-2001; 2001US-276791P.		
PR	16-MAR-2001; 2001US-276888P.		
PR	06-APR-2001; 2001US-281922P.		
PR	24-APR-2001; 2001US-286214P.		
PR	30-APR-2001; 2001US-0847046.		
PR	04-MAY-2001; 2001US-288589P.		
PA			
PA	(EOSB-) EOS BIOTECHNOLOGY INC.		
XX			
Pi	Gish KC, Mack DH, Walsen KE, Afar D, Hevezl P;		
XX			
DR	WPI, 2002-471335/50.		
DR	P-PsDB; ABG61805.		
XX			
PT			
PT	Detecting a prostate cancer-associated transcript in a cell in a		
PT	patient, useful for diagnosing prostate cancer (PC) or screening		
PT	modulators of PC, by determining if prostate cancer-associated genes		
PT	are expressed in a prostate tissue		
PS			
PS	Claim 22; Page 305; 436pp; English.		
CC			
CC	The present invention relates to methods of detecting a prostate		
CC	cancer-associated transcript in a cell from a patient. The method		
CC	comprises contacting a biological sample from the patient with		
CC	prostate cancer-associated polynucleotides (designated PC genes) that		
CC	selectively hybridize to a sequence that is at least 80% identical		
CC	to them. The prostate cancer-associated polynucleotide sequences		
CC	are differentially expressed in prostate tumour tissue or in		
CC	prostate cancer and are derived from the tissues of various		
CC	organisms such as humans or other mammals (e.g. mice, sheep and dogs)		
CC	The methods of the invention are useful for diagnosing and treating		
CC	prostate cancer in mammals. The prostate cancer-associated genes are		
CC	useful for diagnosing or treating prostate cancer, as well as for		



CC Identifying modulators of prostate cancer or agents that inhibit  
 CC prostate cancer. The nucleic acid sequences are particularly useful  
 CC in gene therapy, as a vaccine or in antisense applications.  
 CC AKR92115-AKR92263 represent prostate cancer-associated polynucleotide  
 CC sequences.

XX Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 other;  
 SQ

Query Match 87.6%; Score 754.2; DB 24; Length 1140;  
 Best Local Similarity 98.3%; Pred. No. 1,8e-145;  
 Matches 762; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

QY 87 TTTGTCAGAGCATGGAGATCAGGAGCTGAGATTGTTGATCATATCATATGTTGCT 146
DB 77 TCTCCTGGAAACCAAGCAATGAGGAGCTGAGTTGTTGATCATATCATATGTTGCT 136
QY 147 GGTATGATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 136
DB 137 GGTATGATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 196
QY 207 ACGGTCCTTCATCAGGCGGACAGCCAGGAGGAGGAGAGAGAAAGATGCGCTGCTCAGA 256
DB 197 ACGGTCCTTCATCAGGCGGACAGCCAGGAGGAGGAGAGAGAAAGATGCGCTGCTCAGA 256
QY 267 AGGATGCTGTTGCGCCCTCGAGAGCAGATGTCAGGAGGAGGAGGAGGAGGAGGAGG 326
DB 257 AGGATGCTGTTGCGCCCTCGAGAGCAGATGTCAGGAGGAGGAGGAGGAGGAGGAGG 316
QY 327 CTACGCGCCGCTGCGCCAGCCAGCCGCTGCGCCGCTGCGCCGCTGCGCCGCTGCGCC 386
DB 317 CTACGCGCCGCTGCGCCAGCCAGCCGCTGCGCCGCTGCGCCGCTGCGCCGCTGCGCC 376
QY 387 CTTCACGCGCTGCGCCAGCCAGCCAGCCGCTGCGCCGCTGCGCCGCTGCGCCGCTGCGCC 446
DB 377 CTTCACGCGCTGCGCCAGCCAGCCAGCCGCTGCGCCGCTGCGCCGCTGCGCCGCTGCGCC 436
QY 447 CATCTGCTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 506
DB 437 CATCTGCTGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 496
QY 507 TCGGAGCCCGGAGCAGCAGCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGA 566
DB 497 TCGGAGCCCGGAGCAGCAGCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGA 556
QY 567 AACCATCTTGCAGAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 626
DB 557 AACCATCTTGCAGAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 616
QY 627 CAGTAACTCGGGGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 686
DB 617 CAGTAACTCGGGGATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 676
QY 687 GCCCACTACAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 746
DB 677 GCCCACTACAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 736
QY 747 CAGTGGGCGGCTCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 806
DB 737 CAGTGGGCGGCTCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 796
QY 807 AGAGAGGCGAGCATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 861
DB 797 AGAGAGGCGAGCATGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 851

```

DE cDNA clone encoding a human TANGO 261 polypeptide.

XX TANGO 266; TANGO 216; TANGO 261; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic cancer-associated disease; atelectasis; pulmonary congestion;  
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; latrogenic disease; inflammation; meningitis;  
 KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 XX CDS 6..764  
 XX FT /tag= a  
 XX FT /product= "TANGO 261"

XX W0200052022-A1.  
 XX PD 08-SEP-2000.  
 XX PE 01-MAR-2000; 2000MO-US05226.  
 XX PR 01-MAR-1999; 9905-0122458.  
 XX PA (MILL-) MILLENNIUM PHARM INC.  
 XX PI Barnes TM, Holzman DA, Sharp JD, Fraser CC;  
 XX WPI: 2000-579269/54.  
 XX DR P-PSDB; AAB18461.

PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX  
 XX PS Disclosure; Page -: 175pp; English.

CC AAA75163-65 encode human TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferation, modulate cellular interactions,  
 CC modulate the proliferation, differentiation, and/or function of cells,  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, treat bone marrow, blood  
 CC pulmonary congestion or edema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, atelectasis, bronchial  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral edema,  
 CC hydrocephalus, brain herniations, latrogenic disease, inflammation,  
 CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.

SQ Sequence 969 BP; 211 A; 317 C; 293 G; 148 T; 0 other;  
 Query Match 87.5%; Score 753.6; DB 21; Length 969;  
 Best Local Similarity 99.5%; Pred. No. 2.3e-145;  
 Matches 756; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 102 GGAGATCAGGAGCTGGAGTTGTTGATGATCATCATCATGTTGTTGATGATGATGATGATGAT 161





AC	AAA75165;
XX	
DT	15-JAN-2001
XX	(first entry)

TANGO 265; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
cellular proliferation; cellular differentiation

von Willebrand factor-associated disorder; atherosclerosis; cellular adhesion; hematopoietic associated disease; atelectasis; cell trafficking; cancer; oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis; intestinal disorder; spleen associated disease; renal disorder; cardiovascular disorder; ischemic heart disease; hydrocephalus; brain herniation; latrogenic disease; inflammation; meningoencephalitis; Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease; multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder, ss. Homo sapiens.

Key	Location/Qualifiers
CDS	6..764
	/*tag 9
	/Product= "TANGO 261"
WO200053022-A1.	
08-SEP-2000.	
01-MAR-2000; 2000WO-US05226.	

PR 01-MAR-1999; 9905-0122A58.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX  
DR WPI: 2000-579269/54.  
XX P-PSDB; AABI8463.  
XX  
PT Novel human and murine secreted proteins designated TANGO 216, 261  
PT 262, 266 and 267 useful as modulating agents of cellular processes  
XX e.g. for treating cancer -  
PS Disclosure; Page -, 175pp; English.  
XX

CC AAA75163-65 encode human TANGO 261 proteins. The specification also  
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
CC polypeptides can be used to modulate cellular proliferation, modulate  
CC cellular differentiation and/or modulate cellular adhesion. The  
CC proteins can be used to treat any von Willebrand factor-associated  
CC disorder, regulate extracellular matrix structuring, cellular adhesion  
CC and cell trafficking and/or migration, modulate cellular interactions,  
CC modulate cell adhesion in proliferative disorders, such as cancer,  
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
CC and hematopoietic associated diseases and disorders, osteoarthritis,  
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, asthma  
CC and bronchiectasis, intestinal disorders, spleen associated  
CC diseases, modulate renal disorders, treat cardiovascular disorders such  
CC as ischemic heart disease, modulate the proliferation, differentiation,  
CC and/or function of bone and cartilage cells and to treat bone and/or  
CC cartilage associated diseases or disorder. They may also be used to  
CC treat disorders associated with the ovaries, and cerebral oedema,  
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain  
CC hydrocephalus and encephalitis, and treat hepatic disorders.  
CC note: The present sequence does not appear in the specification; it was  
CC created using information provided.

Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other,

Query Match	87.5%;	Score 753.6;	DB 21;	Length 969
Best Local Similarity	99.5%;	Pred. No. 2.3e-145;		
Matches 756;	Conservative			

[illegible]

Db	2	GGAGATGGCGGAGCGGAGGATTGTTTCAGATCATCATCATCGTGGTGGATGATGGTAT	16
Qy	162	GGTGTGTGTGATCAACGTGTCCTGGCGAGGCCATTCAGAGCTGTGTCTCAGCGTCTTCATTCAG	22
Db	62	GGTGTGTGTGATCAACGTGTCCTGGCGAGGCCATTCAGAGCTGTGTCTCAGCGTCTTCATTCAG	12
Qy	222	CGGGCACAGCCACGAGGGGCGGAGGAGAGAAATGCTCCTGTCTCCAGAAAGATGCTGTGGCC	28
Db	122	CGGGCACAGCCACGAGGGGCGGAGGAGAGAAATGCTCCTGTCTCCAGAGGATGCTGTGGCC	18
Qy	282	CTCGAGAGGCACAGTGTCCAGGAGCAAGGAATCCACAGACCCAGAGTTCAGACCCCGCCCTCG	34
Db	182	CTCGAGAGGCACAGTGTCCAGGAGCAAGGAATCCACAGACCCAGAGTTCAGACCCCGCCCTCG	24
Qy	342	GCCACAGCAACGCTGTGGCGTCCGCGCCCTTGCCGCCAGCGGAGACGCTTCCACACGGTCCA	40
Db	242	GCCACAGCAACGCTGTGGCGTCCGCGCCCTTGCCGCCAGCGGAGACGCTTCCACACGGTCCA	30
Qy	402	GCCACACTTACGTTACCTTCAGCAACGAGATCGACCTGGCCACCCACCATCTCGCTGTAGA	46
Db	302	GCCACACTTACGTTACCTTCAGCAACGAGATCGACCTGGCCACCCACCATCTCGCTGTAGA	36
Qy	462	CGGAGGAGAGCCCGCACCCCTTACGAGGGCCCTTCACACCTTCAGACTTCGAGTTCGGAGACCCGAGCA	52
Db	362	CGGAGGAGAGCCCGCACCCCTTACGAGGGCCCTTCACACCTTCAGACTTCGAGTTCGGAGACCCGAGCA	42

OY	522	GCAGCTGGAACGAAACCGGGAGCTGGTGGCGCACACCCCAACAGAACATCTTGACAG	581		
OY	522	GCAGCTGGAACGAAACCGGGAGCTGGTGGCGCACACCCCAACAGAACATCTTGACAG	581		
Db	422	GCAGCTGGAACGAAACCGGGAGCTGGTGGCGCACACCCCAACAGAACATCTTGACAG	481		
OY	582	TGACCTGATGATAGTGTGCCAGGCTGGGCGGCCCTGGCCCCCAGCAGTAACTCGGGCAT	641		
Db	482	TGACCTGATGATAGTGTGCCAGGCTGGGCGGCCCTGGCCCCCAGCAGTAACTCGGGCAT	541		
OY	642	CAGCGCCACGCTCTACGGCAGCGGGCGGCATGAGAGGGCCGCCCACTACAGCGA	701		
Db	542	CAGCGCCACGCTCTACGGCAGCGGGCGGCATGAGAGGGCCGCCCACTACAGCGA	601		
OY	702	GGTCATTCGGCCACTACCCCGGGGTCCTCCACAGACACGAGAGAGTGGGCCGCCCTC	761		
Db	602	GGTCATTCGGCCACTACCCCGGGGTCCTCCACAGACACGAGAGAGTGGGCCGCCCTC	661		
OY	762	CTTGTGTGAGGGAGCCCGGCTCCACACACACATCGGCCCTTAGAGAGCGCACCAT	821		
Db	662	CTTGTGTGAGGGAGCCCGGCTCCACACACACATCGGCCCTTAGAGAGCGCACCAT	721		
OY	822	CTTGAGCAAAAGAGAGATAAACAGAAAGACACCTCTC	861		
Db	722	CTTGAGCAAAAGAGAGATAAACAGAAAGACACCTCTC	761		
RESULT 7					
ID	AAAA7429	AAAA7429 standard; DNA; 1061 BP.			
AC	AAAA7429;				
DT	20-OCT-2000	(first entry)			
DE	Sequence encoding human neuron-associated protein.				
KM	Neuron associated protein; NEUP; neurological disorder; epilepsy;				
KM	ischemic cerebrovascular disease; stroke; cerebral neoplasm;				
KM	Alzheimer's disease; Pick's disease; Huntington's disease;				
KM	dementia; Parkinson's disease; demyelinating disease; meningitis;				
KM	prion disease; Kuru; Creutzfeldt-Jakob disease; neurofibromatosis;				
KM	cerebral palsy; muscular dystrophy; central nervous system; CNS;				
KM	peripheral nervous system; PNS; myopathy; schizophrenia;				
KM	actinic keratosis; arteriosclerosis; atherosclerosis; burnsitis;				
KM	cirrhosis; hepatitis; mixed connective tissue disease; MCTD;				
KM	myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;				
KM	autoimmune disease; inflammation; acquired immunodeficiency syndrome;				
KM	AIDS; Addison's disease; adult respiratory distress syndrome;				
KM	allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;				
KM	Werner syndrome; trauma; human; ds.				
OS	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	101..859			
FT		/*tag= a			
FT		/product= Neuron associated protein			
PN	WO200034477-A2.				
PD	15-JUN-2000.				
PF	10-DEC-1999;	99WO-US30408.			
XX					
XX	11-DEC-1998;	98US-0210083.			
PR	11-DEC-1998;	98US-0212456.			
PR	09-FEB-1999;	99US-0119365.			
PR	16-MAR-1999;	99US-0124687.			
XX					
PA	(INCY-1)	INCYTE PHARM INC.			
XX					
PI	Tang YT, Yue H, Baughn MR, Hillman JT, Lal P, Au-Young J, Yang J;				
PI	Lu DM, Azimzal Y;				

XX		WPI: 2000-423423/36.
DR	P-PSDB; AAB01388.	
PT	New human neuron-associated proteins and polynucleotides encoding them,	
PR	useful for diagnosis, treatment and prevention of cell proliferative	
PT	disorders including cancer, neuronal and neurological disorders	
PS	Claim 9; Page 136; 145pp; English.	
CC	Human neuron-associated proteins (NEUP) can be used for for	
CC	treating or preventing a disorder associated with decreased	
CC	expression or activity of NEUP. Antagonists of NEUP are useful for	
CC	treating or preventing disorder associated with increased expression	
CC	or activity of NEUP. NEUP or their fragments or derivatives are	
CC	useful for treating neurological disorder such as epilepsy, ischemic	
CC	cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's	
CC	disease, Pick's disease, Huntington's disease, dementia and	
CC	Parkinson's disease. NEUPs are also useful for treating other	
CC	demyelinating diseases, bacterial and viral meningitis, prion	
CC	diseases including kuru, Creutzfeldt-Jakob disease, nutritional and	
CC	metabolic diseases of the nervous system, neurofibromatosis, other	
CC	developmental disorders of the central nervous system, cerebral	
CC	palsy, neuromuscular disorders, autonomic nervous system disorders,	
CC	cranial nerve disorders, spinal cord diseases, muscular dystrophy and	
CC	other neuromuscular disorders, peripheral nervous system disorders,	
CC	inherited, metabolic, endocrine, and toxic myopathies, mental	
CC	disorders including mood, anxiety and schizophrenic disorders, a cell	
CC	proliferative disorder such as actinic keratosis, arteriosclerosis,	
CC	atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective	
CC	tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal	
CC	hemoglobinuria, cancers of the adrenal gland, bladder, bone,	
CC	bone marrow, brain, breast, cervix, and an autoimmune/inflammatory	
CC	disorder such as acquired immunodeficiency syndrome (AIDS), Addison's	
CC	disease, adult respiratory distress syndrome, allergies, ankylosing	
CC	spondylitis, amyloidosis, anemia, asthma, Werner syndrome,	
CC	complications of cancer, hemodialysis, and extracorporeal circulation,	
CC	viral, bacterial, fungal parasitic, protozoal, and helminthic	
CC	infections, and trauma. This sequence was given the Incyte ID no.	
CC	187128BCB1.	
XX		
SQ	Sequence 1061 BP: 225 A; 342 C; 326 G; 168 T; 0 other:	
	Query Match            87.4%; Score 752.6; DB 21; Length 1061;	
	Best Local Similarity   98.2%; Pred. No. 3.7e-145;	
	Matches 761; Conservative     0; Mismatches 14; Indels     0; Gaps       0;	
OY	87 TTTGTTCCAGACCATGGAGAATFCAGCAGACTGTGAGATTGTCAGATCATATCATCGTGT 146	
Dd	82 TCTCTGCCAAACCAGGCATAATGGCCGAGCTGGAGTGTTGTCAGATCATATCATCTGTGT 141	
OY	147 GGTCATGATGGTAGTGTGGTGGTGAATCACTGACGCGCTGACCACTACAACCTGCTTC 206	
Ddb	142 GGTGATGATGGTAGTGTGGTGGTATATCACTGCGCTGCGAGCACCACTAACACTGTCTCG 201	
OY	207 ACGGTCTTCATCAGCCGCGACACGCCAGCGGCGGAGAGAGAAGATGCCCTGTCTCACA 266	
Dd	202 AGCGTCTTCATCAGCCGCGACACGCCAGCGGCGGAGAGAGAAGATGCCCTGTCTCACA 261	
OY	267 AGGATGCGGTGGCCCTCGAGAGACACACTGTCAGGACCAACGGAATCCCAGAGCCGAGGT 326	
Dd	262 AGGATGCGGTGGCCCTCGAGAGACACACTGTCAGGACCAACGGAATCCCAGAGCCGAGGT 321	
OY	327 CTAGCGCCCCCGCGCCACCGACCGCTGGCCGTGCCGCCCTTCGCCAGCGGAGCG 386	
Dd	322 CTAGCGCCCCCGCGCCACCGACCGCTGGCCGTGCCGCCCTTCGCCAGCGGAGCG 381	
OY	387 GTTCACGCGCTTCAGGCCACCTATTCGTAATCGGACAGCGAGATTCGACATCGCACATCC 446	
Dd	382 GTTCACGCGCTTCAGGCCACCTATTCGTAATCGGACAGCGAGATTCGACATCGCGCCAC 441	
OY	447 CATTCGCTGTCAGACGGGAGAGAGCCCCACCTATCACAGGAGCCCTTCGACCTCTCAGCT 506	





Accession	Gene	Species	Length (bp)	Location/Qualifiers
DB 310	--TGGAAATGCCGGAAGCCACAGAGTCTATATGCCCCCGCTCGGAGCCACACTGACCGACTGCTGTG	Human	367	
QY 364	CGGCCCTTCCGCCAGCGGGAGGCGTTCACACGGCTTCCAGCCACCTATCGTATACCTTCAG	Human	423	
DB 368	CCCCCTTATCATCAGCGG-----ACGGCAATTCACACCCACCTACCCCTTACTCTCAG	Human	418	
QY 424	CACGAGATGACCTGTGCACACCCACCATCTCTGCTGTGTAGACGGGAGAGCCCCACCTTAC	Human	483	
DB 419	CAGGAATATTCCTGTGCACACCCACCATCTCTCATCTGTCTGATGTGGGAGAGGCCCTTAC	Human	478	
QY 484	CAGGGCCCTGTGACCCCTTGCAGCTTGGGAGACCCGAGAGAGACAGTGGAACTGAACGGGAG	Human	543	
DB 479	CAGGGCCCTGTGACCCCTTGCAGCTTGGGAGACCCGAGAGAGACAGTGGAACTGAACGGGAG	Human	538	
QY 544	TGGTGGCCGCGACCCCGCCAAACAGACCATCTTGCAGAGTACCTGATGTGATGTGGCAG	Human	603	
DB 539	TCTGTGGCCGCGACCCCGCCAAACAGACCATCTTGCAGAGTACCTTATACAGACACCATG	Human	598	
QY 604	CTGGGCGGCCCTTGCAGCCCGCCAGCAGTAATCTGGGGATACAGCGCCACGTGTACGGCAGC	Human	663	
DB 599	CTGGGCGGCCCTTGCAGCCCGCCAGCAGTAATCTGGGGATACAGCGCCACGTGTACGGCAGC	Human	658	
QY 664	GGCGGGCGCATGTGAGGGGGCGCCGCCACCTTACAGGAGGTATCTGGCCACTTACCGCTC	Human	723	
DB 659	GGTGGGGCGCATGTGAGGGGGCGCCGCCACCTTACAGGAGGTATCTGGCCACTTACCGCTC	Human	718	
QY 724	TCTCTCTTCCAGCAGCAGCAGCAGCAGTGTGGGCGCCCTCTCTGTGTGAGAGGAGACCGGCTC	Human	783	
DB 719	TCTCTCTTCCAGCAGCAGCAGCAGTGTGGGCGCCCTCTCTGTGTGAGAGGAGGAGCGGCTC	Human	778	
QY 784	CACGACACACATCTGCGCGCCCTTAAGAGC 813	Human		
DB 779	CATCACTCGCACATCTGCCCACTGAGAGAC 808	Human		
RESULT 11				
AAAT75152				
ID	AAAT75152 standard; cdna; 1713 BP.			
XX	AAAT75152;			
AC				
DT	15-JAN-2001 (first entry)			
DE	CDNA encoding a murine TANGO 261 polypeptide.			
XX				
KW	TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;			
KW	cellular proliferation; cellular differentiation; cellular adhesion;			
KW	von Willebrand factor-associated disorder; cell trafficking; cancer;			
KW	hematopoietic associated disease; atelectasis; pulmonary congestion;			
KW	oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;			
KW	intestinal disorder; spleen associated disease; renal disorder;			
KW	cardiovascular disorder; ischemic heart disease; hydrocephalus;			
KW	brain herniation; iatrogenic disease; inflammation; meningitis;			
KW	Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;			
KW	multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.			
OS				
XX				
XX	Mus sp.			
FH				
FT	Key	Location/Qualifiers		
FT	CDS	2..655		
FT		/*tag= a		
FT		/product= "TANGO 261"		
XX				
PN	WO200052022-A1.			
XX				
PD	08-SEP-2000.			
XX				
PF	01-MAR-2000; 2000WO-US05226.			
XX				
PR	01-MAR-1999; 99US-0122458.			
XX				
PA	(MILL-) MILLENNIUM PHARM INC.			

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 XX WPI: 2000-579269/54.  
 DR P-PSDB: AAB18450.  
 XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 261, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX Claim 2; Fig 6A-B; 175pp; English.  
 PS The present sequence encodes a murine TANGO 261 polypeptide. The  
 CC specification also describes TANGO 266, TANGO 216, TANGO 262, and  
 CC TANGO 267. The TANGO polypeptides can be used to modulate cellular  
 CC proliferation, modulate cellular differentiation and/or modulate  
 CC cellular adhesion. The proteins can be used to treat any von Willebrand  
 CC factor-associated disorder, regulate extracellular matrix structuring,  
 CC cellular adhesion, and cell trafficking and/or migration, modulate  
 CC cellular interactions, modulate cell adhesion in proliferative  
 CC disorders, such as cancer, modulate the proliferation, differentiation,  
 CC and/or function of cells that appear in the bone marrow, and leukocytes,  
 CC treat bone marrow, blood and hematopoietic associated diseases and  
 CC disorders, atelectasis, pulmonary congestion or edema, emphysema,  
 CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
 CC disorders, spleen associated diseases, modulate renal disorders, treat  
 CC cardiovascular disorders such as ischemic heart disease, modulate the  
 CC proliferation, differentiation, and/or function of bone and cartilage  
 CC cells and to treat bone and/or cartilage associated diseases or  
 CC disorder. They may also be used to treat disorders associated with the  
 CC ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic  
 CC disease, inflammations, bacterial and viral meningitis, Alzheimer's  
 CC disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,  
 CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
 CC disorders.  
 XX  
 XX Sequence 1713 BP; 506 A; 438 C; 400 G; 369 T; 0 other;  
 S0  
 Query Match 53.7%; Score 462.2; DB 21; Length 1713;  
 Best Local Similarity 83.8%; Pred. No. 1.2e-85;  
 Matches 569; Conservative 0; Mismatches 83; Indels 27; Gaps 3;

OY 663 CGGGGGCGCATGAGGGGGCCGCCACCTACAGCAGAGGTCATGGCCATACCCGGG 722  
 DB 469 CGGTGGGCGCATGAGGGGGCCGCCACCTACAGCAGAGGTCATGGCCATACCCGGG 528  
 OY 723 GTCTCTCTTCAGACACCGACAGAGCAGTGGGGCCCTCTCTTGTGTGAGGGAGCCGGCT 782  
 DB 529 CTCCTCTTCAGACACCGACAGCAGAGGTCATGGCCCTCTCTTGTGTGAGGGAGCCGGCT 588  
 OY 783 CCACCACACACATCGGGCCCTTAGAGAGCGGACCATCTGGAGCAAGAAGAGATTA 842  
 DB 589 CCATCTACTCGCACATTTGCCCATCTGA-----GAACAAGAGAAAGAGAA 633  
 OY 843 ACAGAAAGACACCTCTC 861  
 DB 634 ACAGAAAGTACCCCTC 652

RESULT 12  
 AAA75166  
 ID AAA75166 standard; cDNA: 1713 BP.  
 XX  
 AC AAA75166;  
 DT 15-JAN-2001 (first entry)  
 XX  
 DE cDNA clone encoding a murine TANGO 261 polypeptide.  
 XX  
 KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 2..655  
 FT /tag= a  
 FT /product= "TANGO 261"  
 XX  
 XX WO200052022-A1.  
 XX 08-SEP-2000.  
 XX 01-MAR-2000; 2000WO-US05226.  
 XX 01-MAR-1999; 99US-0122458.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PT Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 DR WPI: 2000-579269/54.  
 DR P-PSDB: AAB18464.  
 PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 261, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX  
 PS Disclosure: Page -: 175pp; English.  
 CC AAA75166-68 encode murine TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,



CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, treat bone marrow, blood  
 CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral oedema,  
 CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 CC bacterial and viral meningitis, Alzheimer's disease, Parkinson's  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders,  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.

Sequence 1713 BP; 505 A; 439 C; 400 G; 369 T; 0 other;

Query Match 53.58; Score 460.6; DB 21; Length 1713;  
 Best Local Similarity 83.78; Pred. No. 2,6e-85;  
 Matches 568; Conservative 0; Mismatches 84; Indels 27; Gaps 3;

QY 183 GCTAGCCACTACAAAGCTGTCTCAGCGTCCCTCATCAGCCGACAGCCAGGCGGAG 242  
 Db 1 GCTAGCCACTACAAAGCTGTCTCAGCGTCCCTCATCAGCCGACAGCCAGGCGGAG 242  
 QY 243 GAGAGAGATGCTGCTGCTCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302  
 Db 61 GAGAGAGATGCTGCTGCTCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302  
 QY 303 CAACGAGATCCAGAGCCGACAGTCTACGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 362  
 Db 121 --TGAATGCGGAGCCACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362  
 QY 363 GCGGCGCTTGGCGAGCGGAGCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAG 422  
 Db 178 GCGGCGCTTGGCGAGCGGAGCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAG 422  
 QY 423 GCACGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482  
 Db 229 GCACGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482  
 QY 483 CCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542  
 Db 289 CCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542  
 QY 543 GTCGCTGCGGCGACCCCAACAGAACCATCTTGCAGAGTGCATGATGATGATGATGATGAT 602  
 Db 349 ATCTGTGCGCGACCCCTTAACGGACCATCTTGCAGAGTGCATGATGATGATGATGATGAT 602  
 QY 603 GCTGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662  
 Db 409 GCTGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662  
 QY 663 GCGGCGCGCGATGAGGCG 722  
 Db 469 GCGGCGCGCGATGAGGCG 722  
 QY 723 GTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782  
 Db 529 GTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782  
 QY 783 CCACGACACAGATGCG 842  
 Db 589 CCATCTCTCTGCGACG 842  
 QY 843 ACAGAAAGGACACCTCTTC 861  
 Db 634 ACAGAAAGGACACCTCTTC 861  
 634 ACAGAAAGGACACCTCTTC 652

RESULT 13  
 AAA75167  
 ID AAA75167 standard; CDNA: 1713 BP.  
 AAA75167;  
 15-JAN-2001 (first entry)  
 CDNA clone encoding a murine TANGO 261 polypeptide.  
 TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 cellular proliferation; cellular differentiation; cellular adhesion;  
 von Willibrand factor-associated disorder; cell trafficking; cancer;  
 hematopoietic associated disease; atelectasis; pulmonary congestion;  
 oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 intestinal disorder; spleen associated disease; renal disorder;  
 cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 brain herniation; iatrogenic disease; inflammation; meningitis;  
 Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
 multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
 Mus sp.  
 Key Location/Qualifiers  
 CDS 2..655  
 /tag= a  
 /product= "TANGO 261"  
 WO200052022-A1.  
 08-SEP-2000.  
 01-MAR-2000; 2000WO-US05226.  
 01-MAR-1999; 99US-0122458.  
 (MILL.) MILLENNIUM PHARM INC.  
 Barnes TM, Holzman DA, Sharp JD, Fraser CC;  
 WPI: 2000-579269/54.  
 P-PSDB: AAB18465.  
 Novel human and murine secreted proteins designated TANGO 216, 261,  
 262, 266 and 267 useful as modulating agents of cellular processes,  
 e.g. for treating cancer.

Disclosure: Page -: 175pp; English.

AAA75166-68 encode murine TANGO 261 proteins. The specification also  
 describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 polypeptides can be used to modulate cellular proliferation, modulate  
 cellular differentiation and/or modulate cellular adhesion. The  
 proteins can be used to treat any von Willibrand factor-associated  
 disorder, regulate extracellular matrix structuring, cellular adhesion,  
 and cell trafficking and/or migration, modulate cellular interactions,  
 modulate cell adhesion in proliferative disorders, such as cancer,  
 that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 and hematopoietic associated diseases and disorders, treat bone marrow, blood  
 pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 asthma and bronchiectasis, intestinal disorders, spleen associated  
 diseases, modulate renal disorders, treat cardiovascular disorders such  
 as ischemic heart disease, modulate the proliferation, differentiation,  
 and/or function of bone and cartilage cells and to treat bone and/or  
 cartilage associated diseases or disorder. They may also be used to  
 treat disorders associated with the ovaries, and cerebral oedema,  
 hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 bacterial and viral meningitis, Alzheimer's disease, Parkinson's  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders,  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.





Db 121 ---TGGAATGCCGACACAGTCTATGCCCCGCTTGGCCCACTAGACCGACTCGCTGT 177  
 QY 363 GCCGCCCTTGGCCAGGGGAGCGCTTCCACCGCTTCCAGCCGACCTATCCGTACCTGCA 422  
 Db 178 GCCGCCCTTCAATCCAGGG-----AGCCGATTCACCAACCCACCTACCCCTACTGCA 228  
 QY 423 GCACGAGATGCACTGGCCACCAACATCTGCTGTGAGAGGGGAGAGGCCCCACCTTA 482  
 Db 229 GCACGAAATTCGCTGCGACCAACATCTGCTGTGATGGAGAGAGGCCCCACCTTA 288  
 QY 483 CCAGGCGCCCTGCAACCTTCCAGCTTGGGAGCCCGAGACGACCTGGAATGACCGGGA 542  
 Db 289 CCAGGCGCCCTGCAACCTTCCAGCTTGGGAGCCCGAGACGACCTGGAATGACCGGGA 348  
 QY 543 GTGGGTGCGGACCCCAAGAACATCTTGCAGCTGACGCTGATGATGATGATGATG 602  
 Db 349 ATCTGTGCGGACCCCTTACCGGACCATCTTGCAGCTGATGATGATGATGATGATG 408  
 QY 603 GCTGGGCGGCGCCCTTGGCCCGCCAGCTAAGTGGGATGAGCGGACGCTGATGAGGAG 662  
 Db 409 GCTGGGCGGCGCCCTTGGCCCGCCAGCTAAGTGGGATGAGCGGACGCTGATGAGGAG 468  
 QY 663 CGGCGGCGGCGGAGGAGCGCCCGCCAGCTAAGTGGGATGAGCGGACGCTGATGAGGAG 722  
 Db 469 CGGCGGCGGCGGAGGAGCGCCCGCCAGCTAAGTGGGATGAGCGGACGCTGATGAGGAG 528  
 QY 723 GTCTCTCTTCCAGCAACAGACAGATGGGCGCCCTTCTGTGAGGAGGAGCGGAGT 782  
 Db 529 GTCTCTCTTCCAGCAACAGACAGATGGGCGCCCTTCTGTGAGGAGGAGCGGAGT 588  
 QY 783 CCAACACACACACATGCGGCGCCCTTGAAGAGCGGACCATCTGTGAGGAGGAGGATTA 842  
 Db 589 CCATCATCTGCGACATTTGCCCTCCCTGGA-----GAACAAGGAGGAGGAGAA 633  
 QY 843 ACAGAAAGGACACCCCTCTC 861  
 Db 634 ACAGAAAGTACCCCTCTC 652

RESULT 15  
 ABK12143/C  
 ID ABK12143 standard; cDNA; 693 BP.  
 AC ABK12143;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Human MIVR-1 homologous sequence #1.  
 XX  
 KW Human; ss; MIVR-1; Mechanically Induced Vascular Receptor 1;  
 KW cytoskeletal; cardiant; cerebroprotective; antiarteriosclerotic;  
 KW cardiac cell; anti-apoptotic; vascular endothelial cell;  
 KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;  
 KW heart failure; A1761441.1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200216416-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 21-AUG-2001; 2001WO-US26089.  
 XX  
 PR 22-AUG-2000; 2000US-227159P.  
 XX  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 PA (PFIZ ) PFIZER INC.  
 PI Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;  
 XX WPI; 2002-280912/32.  
 XX  
 PT Novel nucleic acid molecule encoding Mechanically Induced Vascular

PT Receptor-1 polypeptide, useful for treating cardiovascular diseases -  
 XX  
 PS Disclosure; Page 101; 105pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC Mechanically induced Vascular Receptor (MIVR)-1 polypeptide having  
 CC cardiac cell anti-apoptotic activity and fragments of it provided  
 CC they are not identical to Genbank sequences A1761441.1, A1594390,  
 CC NM 004338 and A0177461. Also included are expression vectors, host  
 CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
 CC MIVR-1, a candidate agent, where the molecule is a nucleic acid molecule  
 CC comprising MIVR-1, IEX-1, VDR-1, BRC-2 and TIS-110 or its  
 CC expression product, determining if the anti-apoptotic activity is  
 CC modulated and thereby identifying a modulator. The cardiac cell anti-  
 CC apoptotic molecules and nucleic acids of the invention are useful for  
 CC treating, diagnosing and monitoring progression of such diseases and  
 CC disorders as characterized by increased apoptotic cell-death of vascular  
 CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
 CC stroke, arteriosclerosis and heart failure. The present sequence  
 CC is one of the four Genbank sequences (A1761441.1) which are homologous to  
 CC the cDNA for human MIVR-1 and which are specifically disclaimed.  
 CC  
 XX  
 SQ Sequence 693 BP; 101 A; 205 C; 237 G; 149 T; 1 other;

Query Match 51.8%; Score 445.6; DB 24; Length 693;  
 Best Local Similarity 97.2%; Pred. No. 2,7e-82;  
 Matches 485; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 364 CGGCCCTTGGCCAGGGAGGCGCTTCCAGCGCTTCCAGCCGCTATTCGTACCTGAG 423  
 Db 693 CGGCCCTTGGCCAGGGAGGCGCTTCCAGCGCTTCCAGCCGCTATTCGTACCTGAG 634  
 QY 424 CACGAGATGCACTTGCACCCACCATCTGCTGTGAGAGGAGGAGGCGCCACCTTAC 483  
 Db 633 CACGAGATGCA-CTGCGCGCCACCATCTGCTGTGAGAGGAGGAGGCGCCACCTTAC 577  
 QY 484 CAGGCGCCCTGACCTTCCAGCTTCCAGCGGAGCGAGCAGTGTGAATGAACCGGAG 543  
 Db 576 CAGGCGCCCTGACCTTCCAGCTTCCAGCGGAGCGAGCAGTGTGAATGAACCGGAG 517  
 QY 544 TCGGTGCGGCGACCCCAACAGAACCATCTTGCAGAGTACCTGTATGATGTCAGAG 603  
 Db 516 TCGGTGCGGCGACCCCAACAGAACCATCTTGCAGAGTACCTGTATGATGTCAGAG 457  
 QY 604 CTGGCGGCGCCCTGCGCCCGCCAGCACTTGGGATCAGCGCCGCTGTAAGGAGC 663  
 Db 456 CTGGCGGCGCCCTGCGCCCGCCAGCACTTGGGATCAGCGCCGCTGTAAGGAGC 397  
 QY 664 GCGGCGGCGATGAGAGGCGCGCCCGCCAGCTAGCGAGTATGCGGACATACCGGGG 723  
 Db 396 GCGGCGGCGATGAGAGGCGCGCCCGCCAGCTAGCGAGTATGCGGACATACCGGGG 337  
 QY 724 TCCCTCTTCCAGACGAGAGAGCACTGGGCGCGCCCTCTTCTGTA-GGGAGCCGAGT 782  
 Db 336 TCCCTCTTCCAGACGAGAGCACTGGGCGCGCCCTCTTCTGTAAGGAGGAGGAGT 277  
 QY 783 CCAACACACACATGCGCCCGCTGAGAGCGCAGCATCTGAGAGCAAGAGAGATTA 842  
 Db 276 CCGCCACACACATGCGCCCGCTGAGAGCGCAGCATCTGAGAGCAAGAGAGATTA 217  
 QY 843 ACAGAAAGGACACCCCTCTC 861  
 Db 216 ACAGAAAGGACACCCCTCTC 198

Search completed: March 17, 2003, 12:28:57  
 Job time : 238.386 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 12:18:54 ; Search time 41.4322 Seconds  
(without alignments)  
6373.035 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861

Sequence: 1 atgacacgccttgatg9ggc.....aacagaaagcaccctctc 861

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCtUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/Backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352.2	40.9	921	US-09-091-952A-7	Sequence 7, Appli
2	352.2	40.9	8065	US-09-091-952A-6	Sequence 6, Appli
3	284.2	33.0	867	US-09-091-952A-8	Sequence 8, Appli
4	47.8	5.6	8438	US-07-945-283-1	Sequence 1, Appli
5	46.2	5.4	15378	US-08-785-420-1	Sequence 1, Appli
6	45.4	5.3	424	US-08-476-176B-7	Sequence 7, Appli
7	45.4	5.3	424	US-08-127-721A-7	Sequence 7, Appli
8	45.4	5.3	424	US-08-485-246A-7	Sequence 7, Appli
9	44	5.1	1293	US-08-924-440-1	Sequence 1, Appli
10	44	5.1	44377	US-08-804-227C-7	Sequence 7, Appli
11	44	5.1	44377	US-08-804-198-1	Sequence 1, Appli
12	43.8	5.1	424	US-08-476-176B-5	Sequence 5, Appli
13	43.8	5.1	424	US-08-476-176B-9	Sequence 9, Appli
14	43.8	5.1	424	US-08-127-721A-5	Sequence 5, Appli
15	43.8	5.1	424	US-08-127-721A-9	Sequence 9, Appli
16	43.8	5.1	424	US-08-485-246A-5	Sequence 5, Appli
17	43.8	5.1	424	US-08-485-246A-9	Sequence 9, Appli
18	43.8	5.1	6854	US-09-194-905-7	Sequence 7, Appli
19	43.2	5.0	3705	US-08-474-379C-64	Sequence 64, Appli
20	43.2	5.0	3705	US-09-146-249A-64	Sequence 64, Appli
21	43.2	5.0	3705	US-08-206-188B-64	Sequence 64, Appli
22	43.2	5.0	152331	US-09-128-155-16	Sequence 16, Appli
23	42.8	5.0	4183	US-08-996-083-2	Sequence 2, Appli
24	42.8	5.0	4183	US-09-429-516-2	Sequence 2, Appli
25	42.2	4.9	1085	US-09-509-902A-3	Sequence 3, Appli
26	42.2	4.9	2793	US-08-209-747-1	Sequence 1, Appli
27	42.2	4.9	2793	US-08-458-298-1	Sequence 1, Appli

28	42.2	4.9	18596	US-09-318-448-11	Sequence 11, Appli
29	41.6	4.8	30001	US-08-125-468-1	Sequence 1, Appli
30	41.6	4.8	30001	US-08-474-933-1	Sequence 1, Appli
31	41.4	4.8	693	US-09-402-515A-17	Sequence 17, Appli
32	41.2	4.8	377	US-08-332-766A-1	Sequence 1, Appli
33	41	4.8	888	US-08-765-907A-2	Sequence 2, Appli
34	41	4.8	2277	US-08-676-967-5	Sequence 5, Appli
35	41	4.8	2277	US-08-676-974-5	Sequence 5, Appli
36	41	4.8	2277	US-09-098-487-5	Sequence 5, Appli
37	41	4.8	2888	US-08-765-907A-1	Sequence 1, Appli
38	40.8	4.7	1743	US-08-665-259-20	Sequence 20, Appli
39	40.8	4.7	1743	US-08-762-500-20	Sequence 20, Appli
40	40.8	4.7	1974	US-08-762-500-78	Sequence 78, Appli
41	40.8	4.7	6803	US-08-665-259-19	Sequence 19, Appli
42	40.8	4.7	6803	US-08-762-500-19	Sequence 19, Appli
43	40.8	4.7	28804	US-08-592-874-1	Sequence 1, Appli
44	40.8	4.7	28804	US-09-096-942-2	Sequence 2, Appli
45	40.8	4.7	28804	US-09-096-867-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-09-091-952A-7  
Sequence 7, Application US/09091952A  
Patent No. 6458532  
GENERAL INFORMATION:  
APPLICANT: Detera-Wadleigh, Sevilla D.  
Gershon, Elliot S.  
Badner, Judith A.  
Goldin, Lynn R.  
Bertellini, Wade H.  
Yoshikawa, Takeo  
Sanders, Alan R.  
Esterling, Lisa E.  
TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,952A  
FILING DATE: 19-Apr-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,278  
FILING DATE: 28-Oct-1996  
APPLICATION NUMBER: PCT/US97/19381  
FILING DATE: 28-Oct-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 015280-297100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: Linear  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1...921  
 OTHER INFORMATION: Clone 22 coding region  
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Query Match 40.9%; Score 352.2; DB 4; Length 921;  
 Best Local Similarity 73.1%; Pred. No. 1.4e-69;  
 Matches 482; Conservative 0; Mismatches 168; Indels 9; Gaps 2;

```

OY 94 CAGAGCATGAGATCATCGAGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGATG 153
DB 166 CCGGGCATCTTCACTCGAGAGCTGGAGTTCCGCCAATCATCATCATCATCGTGGTGGATG 225
OY 154 ATGTGATGTGTGTGTATACGTCCTGCTGAGCCACTACAAAGCTGTGTGACGGTGC 213
DB 226 ACGGTGATGTGTGTATACGTCCTGCTGAGCCACTACAAAGCTGTGTGACGGTGC 285
OY 214 TTCATCAGCGCCGACAGCAGCGGCGGAGAGAGAGAGATCCCTGTCTCAGAGAGATGC 273
DB 286 TTCATCAGCGCCGACAGCAGCGGCGGAGAGAGAGAGATCCCTGTCTCAGAGAGATGC 345
OY 274 CTGTGCGCTTGGAGAGCAGAGTGTCAAGCAAGCAATCCAGCGGCTCTTACGCC 333
DB 346 CTGTGCGCTTGGAGAGCAGAGTGTCAAGCAAGCAATCCAGCGGCTCTTACGCC 399
OY 334 CCGGCTGCGCCGACAGCAGCGGCGGAGAGAGAGATCCCTGTCTCAGAGAGATGC 393
DB 400 GCGGCTGCGCCGACAGCAGCGGCGGAGAGAGAGATCCCTGTCTCAGAGAGATGC 459
OY 394 CCGGCTGCGCCGACAGCAGCGGCGGAGAGAGAGATCCCTGTCTCAGAGAGATGC 453
DB 460 CCGGCTGCGCCGACAGCAGCGGCGGAGAGAGAGATCCCTGTCTCAGAGAGATGC 519
OY 454 CTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 513
DB 520 CTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579
OY 514 CCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 573
DB 580 CCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 639
OY 574 TTGCACTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
DB 640 TTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 699
OY 631 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 690
DB 700 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 759
OY 691 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 749
DB 760 AACTGCGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 818

```

RESULT 2  
 US-09-091-952A-6  
 Sequence 6, Application US/09091952A  
 Patent No. 6458532

GENERAL INFORMATION:  
 APPLICANT: Deterra-Medleigh, Sevilla D.  
 Gershon, Elliot S.  
 Badner, Judith A.  
 Goldin, Lynn R.  
 Berrettini, Made H.  
 Yoshikawa, Takeo  
 Sanders, Alan R.  
 Esterling, Lisa R.

TITLE OF INVENTION: Chromosomal Markers and Diagnostic  
 Tests for Manic-Depressive Illness

NUMBER OF SEQUENCES: 197  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: CA

COUNTRY: USA  
 ZIP: 94111-3834

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 OPERATING SYSTEM: DOS

SOFTWARE: FASTED for Windows Version 2.0  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/091,952A  
 FILING DATE: 19-Apr-1999  
 CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/029,278  
 FILING DATE: 28-OCT-1996

APPLICATION NUMBER: PCT/US97/19381  
 FILING DATE: 28-OCT-1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Smith, Timothy L.  
 REGISTRATION NUMBER: 35,367  
 REFERENCE/DOCKET NUMBER: 015280-297100US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300

TELEX: <unknown>  
 INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 8065 base pairs  
 TYPE: nucleic acid

STRANDEDNESS: single  
 TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
 FEATURE:

NAME/KEY: -  
 LOCATION: 1...8065

OTHER INFORMATION: Clone 22

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 116...1036

OTHER INFORMATION: Clone 22 coding region

FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 452...505

OTHER INFORMATION: alternatively spliced portion

NAME/KEY: misc\_feature  
 LOCATION: 5595...5685

OTHER INFORMATION: amplified region for genotyping

US-09-091-952A-6  
 Query Match 40.9%; Score 352.2; DB 4; Length 8065;  
 Best Local Similarity 73.1%; Pred. No. 2.3e-69;  
 Matches 482; Conservative 0; Mismatches 168; Indels 9; Gaps 2;

```

OY 94 CAGAGCATGAGATCATCGAGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGATG 153
DB 281 CCGGGCATCTTCACTCGAGAGCTGGAGTTCCGCCAATCATCATCATCATCGTGGTGGATG 340
OY 154 ATGTGATGTGTGTGTATACGTCCTGCTGAGCCACTACAAAGCTGTGTGACGGTGC 213
DB 341 ACGGTGATGTGTGTATACGTCCTGCTGAGCCACTACAAAGCTGTGTGACGGTGC 400
OY 214 TTCATCAGCGCCGACAGCAGCGGCGGAGAGAGAGATCCCTGTCTCAGAGAGATGC 273
DB 401 TTCATCAGCGCCGACAGCAGCGGCGGAGAGAGAGATCCCTGTCTCAGAGAGATGC 333
OY 274 CTGTGCGCTTGGAGAGCAGAGTGTCAAGCAAGCAATCCAGCGGCTCTTACGCC 399

```

```
Db 461 CTGTGGCTTACAGACAGCGCCGACCGCGGTGGG-----CGCCTGGAGATCATGCAT 514
Qy 334 CCGGCTGGGCGCCAGCAGCGCTGGCGGTGGCGCCCTTGGCCAGCGGAGCGTTCCAC 393
Db 515 GCGCCGGGCTCAGAGAGGTTCACAGCGCGCTTCATTCACAGAGGATCGTTACGC 574
Qy 394 CGCTTCAGCGCCCTATCCCTACCTGACAGCAGATGACATGCGCCACCACCATCTCG 453
Db 575 CGCTTCAGCGCCCTATCCCTATGATGACAGCAGATGATGATGATGATGATGATGATG 634
Qy 454 CTGTCAAGCGGGAGAGCGCCCGACCCCTACCGAGGCGCCCTGACCTCCAGCTTGGGAC 513
Db 635 CTGTCCAGCGGTGAAGAGCGCTCTTACCGAGGCGCCCTGACCTCCAGCTCCGGGAC 694
Qy 514 CCGGAGAGCAGCTGAGCACTAACCGGAGTGTGGCGCGACCGCCCAAGAGAACATC 573
Db 695 CCGTACAGCAGGATGAGCACTAACCGAGTGTGGAGGCGCCCGACCGAACGACATA 754
Qy 574 TTTCAGAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
Db 755 TTTCAGAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 814
Qy 631 AACTGGGCGATCAGCGCCAGCTGTCTACGCGAGCGCGCGCCATGAGAGGCGCGCGCC 690
Db 815 AACTGGGCGATCAGTGAAGAGCAGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 874
Qy 691 AACTACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
Db 875 ACATACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 933
```

## RESULT 3

US-09-091-952A-8  
Sequence 8, Application US/0901952A  
Patent No. 6458532

## GENERAL INFORMATION:

APPLICANT: Deterra-Wadleigh, Sevilla D.

INVENTOR: Gershon, Elliott S.

INVENTOR: Badner, Judith A.

INVENTOR: Goldin, Lynn R.

INVENTOR: Berretini, Wade H.

INVENTOR: Yoshikawa, Takeo

INVENTOR: Sanders, Alan R.

INVENTOR: Esterling, Lisa E.

TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness

NUMBER OF SEQUENCES: 197

CORRESPONDENCE ADDRESS: Townsend and Townsend and Crew LLP

ADDRESS: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM: MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/091,952A

FILING DATE: 19-Apr-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA: APPLICATION NUMBER: US 60/029, 278

FILING DATE: 28-OCT-1996

APPLICATION NUMBER: PCT/US97/19381

FILING DATE: 28-OCT-1997

ATTORNEY/AGENT INFORMATION: NAME: Smith, Timothy L.

REGISTRATION NUMBER: 35,367

REFERENCE/DOCKET NUMBER: 015280-297100US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 867 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..867  
OTHER INFORMATION: Clone 22 isoform 2 alternatively spliced coding region  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-091-952A-8

Query Match 33.0%; Score 284.2; DB 4; Length 867;  
Best Local Similarity 68.7%; Pred. No. 1.6e-54;  
Matches 453; Conservative 0; Mismatches 143; Indels 63; Gaps 2;

```
Qy 94 CAGAGCATGAGATCAGGAGAGCTGAGTTGTCAGATCATCATCATGTTGTTGATG 153
Db 166 CCGGCGATCTTCAACTGAGAGCTGAGTTGTCAGATCATCATCATGTTGTTGATG 225
Qy 154 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 213
Db 226 ACGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 285
Qy 214 TTTCATGACCGCGGACAGCAGCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 273
Db 286 TTTCATGACCGCGGACAGCAGCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 329
Qy 274 CTGTGGCTTACAGACAGCTGTCTACGCGAGCGCGCGCCATGAGAGGCTTACGCC 333
Db 330 -----GCCGACATCATGCAT 345
Qy 334 CCGGCTGGGCGCCAGCAGCGCTGGCGGTGGCGCCCTTGGCCAGCGGAGCGTTCCAC 393
Db 346 GCGCCGGGCTCAGAGAGGTTCACAGCGCGCTTCATTCACAGAGGATCGCTTACGC 405
Qy 394 CGCTTCAGCGCCCTATCCCTACCTGACAGCAGATGATGATGATGATGATGATGATG 453
Db 406 CGCTTCAGCGCCCTATCCCTATGATGACAGCAGATGATGATGATGATGATGATG 465
Qy 454 CTGTCAAGCGGGAGAGCGCCCGACCCCTACCGAGGCGCCCTGACCTCCAGCTTGGGAC 513
Db 466 CTGTCCAGCGGTGAAGAGCGCACCTCTTACAGAGGCGCTGACCTCCAGCTCCGGGAC 525
Qy 514 CCGGAGAGCAGCTGAGCACTAACCGGAGTGTGGCGCGACCGCCCAAGAGAACATC 573
Db 526 CCGTACAGCAGGATGAGCACTAACCGAGTGTGGAGGCGCCCGACCGAACGACATA 585
Qy 574 TTTCAGAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
Db 586 TTTCAGAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 645
Qy 631 AACTGGGCGATCAGCGCCAGCTGTCTACGCGAGCGCGCGCCATGAGAGGCGCGCGCC 690
Db 646 AACTGGGCGATCAGTGAAGAGCAGCTGACGAGGAGTGAAGAGGAGGAGGAGGAGG 705
Qy 691 AACTACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
Db 706 ACATACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 764
```

## RESULT 4

US-07-945-283-1/c  
Sequence 1, Application US/07945283  
Patent No. 5352596  
GENERAL INFORMATION:



Best Local Similarity 50.7%; Pred. No. 0.17;  
Matches 11; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

OY	325	GTTACGGCCCCGCTGGCCGACACGACCGCTGGCGCGCGCCCTTGACCCAGGGGGAG	384
Db	13329	GTCGGCTTCGGCCGCGGGGGCCACCGCGGTGCTCGCGGTACCTGTGGCCCGTGGGATC	13270
OY	385	CGCTTCCACGCGTTTCCAGCCCACTATCCGTAACGTGACAGCAGAGATGAGCTGCACCC	444
Db	13269	GGGCATGCGCGGACGAGAGCTGGTCAACGCTACCTTTGGGACCGCTCCACGAGGCCGC	13210
OY	445	ACCATCTTGGCTGTCAAGAGGGGAGAGCCCACTCTACAGGGCCCTGCACCTCCAG	504
Db	13209	ACCAAGAGCAGAGCCCAAGAGAGGCGCAGAGCGCTCGCGCCCGGCGCCGCGC	13150
OY	505	CTTGGGAGCCCGCAGACAGCTGGAACTGAAACCGGGAG	543
Db	13149	CCCCGCGTGGCGCAAGCGCCGCCAGAGCAGCGCGGCCAG	13111

```

1      RESULT 6
2      US-08-476-176B-7
3      Sequence 7, Application US/08476176B
4      Patent No. 5958708
5      GENERAL INFORMATION:
6      APPLICANT: Hardman, No. 5958708man
7      APPLICANT: Kolbinger, Frank
8      APPLICANT: Saldanha, Jose
9      TITLE OF INVENTION: Reshaped monoclonal antibodies against an
10     TITLE OF INVENTION: Immunoglobulin isotype
11     NUMBER OF SEQUENCES: 55
12     CORRESPONDENCE ADDRESSES:
13     ADDRESSEE: No. 5958708artis Patent Department
14     STREET: 59 Route 10
15     CITY: East Hanover
16     STATE: New Jersey
17     COUNTRY: USA
18     ZIP: 07936-1080
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: IBM PC compatible
22     OPERATING SYSTEM: PC-DOS/MS-DOS
23     SOFTWARE: Patentln Release #1.0, Version #1.25
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/476,176B
26     FILING DATE:
27     CLASSIFICATION: 435
28     PRIOR APPLICATION DATA:
29     APPLICATION NUMBER: 08/127,721
30     FILING DATE: 27-SEPTEMBER-1993
31     APPLICATION NUMBER: US 07/952, 802
32     FILING DATE: 25-SEPTEMBER-1992
33     ATTORNEY/AGENT INFORMATION:
34     NAME: No. 5958708ak, Henry P.
35     REGISTRATION NUMBER: 33,200
36     REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
37     TELECOMMUNICATION INFORMATION:
38     TELEPHONE: (908) 277-5110
39     TELEFAX: (908) 277-4306
40     INFORMATION FOR SEQ ID NO: 7:
41     SEQUENCE CHARACTERISTICS:
42     LENGTH: 424 base pairs
43     TYPE: nucleic acid
44     STRANDEDNESS: double
45     TOPOLOGY: linear
46     MOLECULE TYPE: DNA (genomic)
47     FEATURE:
48     NAME/KEY: CDS
49     LOCATION: 22..402
50     FEATURE:
51     NAME/KEY: mat_peptide
52     LOCATION: 82..402
53     OTHER INFORMATION: /product="light chain variable

```

OTHER INFORMATION: region C21-L2"  
US-08-476-176B-7

Query Match	5.38;	Score 45.4;	DB 2;	Length 424;
Best Local Similarity	47.48;	Pred. No. 0.11;		
Matches 136;	Conservative 0;	Mismatches 151;	Indels 0;	Gaps 0;

Oy	483	CGAGGGCCCTTCGACACCTTCACGCTTCGGGAGCCCGAGACAGAGCTGGAACTGAACCGGGA	54.2
Db	96	CCAGAGCCCCCGGACACCTTCGAGACCTTCGAGCCCGGGGAGAGGGGCCACCTTCGACTTCGABGC	155
Oy	543	GTCGGTGCGCACACCCCAACAGAACCCATCTTCGACAGTGATGACTGATAGTATAGTGGCAG	60.2
Db	156	CAGCAGAGGACATTCGGACACCACTCCACTGTGTACACAGAGAGCCCGGACAGGCCCCCAG	215
Oy	603	GCTGGGGGGCCCCCTGCCCCCACCACATATACCTGGGGGATACGGCCACGTCATACGGCAG	66.2
Db	216	GCTGCTGATCAAGTATACGCGACGAGAGAGATATAGCGGATATCCCGACAGGTTTCAGGGCAG	275
Oy	663	CGCGGGGGGCGATGAGGGGGCGCGCCCACTTACAGGAGGATCATTCGCGCCACTACCCGGG	72.2
Db	276	CGCGAGCGGACACCGACTTCACCTTCAGACCATATCAGCAGGCTGTAGAGCCCGAGGACACTTCGCCAT	335
Oy	723	GTCCTCCTTCACGACACCAAGCAGACATATGGGGCCGCTCTCTCTGTTGG	76.9
Db	336	GTACTACTGCGCAGCAGAGCGCAGACACTGACCCACCACTTCGGCCAGG	382

```

1      RESULT 7
2      US-08-127-721A-7
3      Sequence 7, Application US/0812721A
4      Patent No. 6066718
5      GENERAL INFORMATION:
6      APPLICANT: Hardman, No. 6066718man
7      APPLICANT: Kolbinger, Frank
8      APPLICANT: Saldanha, Jose
9      TITLE OF INVENTION: Reshaped monoclonal antibodies against an
10     TITLE OF INVENTION: Immunoglobulin isotype
11     NUMBER OF SEQUENCES: 35
12     CORRESPONDENCE ADDRESSES:
13     ADDRESSEE: No. 6066718artis Patent and Trademark Department
14     STREET: 59 Route 10
15     CITY: East Hanover
16     STATE: New Jersey
17     COUNTRY: USA
18     ZIP: 07936-1080
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: IBM PC compatible
22     OPERATING SYSTEM: PC-DOS/MS-DOS
23     SOFTWARE: PatentIn Release #1.0, Version #1.25
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/127,721A
26     FILING DATE: 27-SEPTEMBER-1993
27     CLASSIFICATION: 424
28     PRIOR APPLICATION DATA:
29     APPLICATION NUMBER: US 07/952,802
30     FILING DATE: 25-SEPTEMBER-1992
31     ATTORNEY/AGENT INFORMATION:
32     NAME: No. 6066718ak, Henry P.
33     REGISTRATION NUMBER: 33,200
34     REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP
35     TELECOMMUNICATION INFORMATION:
36     TELEPHONE: (908) 277-5110
37     TELEFAX: (908) 277-4306
38     INFORMATION FOR SEQ ID NO: 7:
39     SEQUENCE CHARACTERISTICS:
40     LENGTH: 424 base pairs
41     TYPE: nucleic acid
42     STRANDEDNESS: double
43     TOPOLOGY: linear
44     MOLECULE TYPE: DNA (genomic)
45     FEATURE:

```

NAME/KEY: CDS  
 LOCATION: 22..402  
 FEATURE:  
 NAME/KEY: mat\_peptide  
 LOCATION: 82..402  
 OTHER INFORMATION: /product="light chain variable"  
 US-08-127-721A-7

Query Match  
 Best Local Similarity 47.4%; Score 45.4; DB 3; Length 424;  
 Matches 136; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 483 CAGGGGCCCCCTGCACCTTCAGCTTCGAGACCCCGAGCAGCAGCTGGAACTGAACCGGGA 542  
 DB 96 CCAGAGCCCCGGCAGCCCTGAGCCCTGAGCCCGGCGAGAGGGCCACCCCTGAGCTGCAGGGC 155  
 QY 543 GTGCGTGGCGCAGACCCCAACAGAACCATCTTGACAGTGAACCTGTATGATGTGCGAG 602  
 DB 156 CAGCCAGACATCGGCACCAACATCTGATGACAGCAGAACGCCGCGCAGGCCCCCGAG 215  
 QY 603 GCGGGGCGGCCCCCTGCCCCCAGCAGTAACTCGGGCATCAGGCGCAGCTGCTAGCGCAG 662  
 DB 216 GCTGCTGATCAAGTACGCGCAGCAGCAGCATCAGCGCATCCCGCAGAGGTTGAGCGCGAG 275  
 QY 663 GCGCGGGCGCATGAGGGGGCGCGCGCCTTACAGCAGGTCATCGGCGCATACCGCGG 722  
 DB 276 CGGCGAGCGGCGACGCTTCAACCTGACCATCAGCAGCTGAGAGCCCGAGAGACTTGGCCAT 335  
 QY 723 GTCCCTCTTCAGCAGCAGCAGCAGTGGGCGCGCCCTCTGCTGG 769  
 DB 336 GTACTACTGCCAGCAGCAGCAGCTGGGCCACCACTTCGGCCAGG 382

## RESULT 8

US-08-485-246A-7  
 Sequence 7, Application US/08485246A  
 Patent No. 6072035  
 GENERAL INFORMATION:

APPLICANT: Hardman, No. 6072035man  
 APPLICANT: Saldanha, Jose  
 TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
 NUMBER OF SEQUENCES: 55  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 6072035artis Patent Department  
 STREET: 59 Route 10  
 CITY: East Hanover  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07936-1080

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/485,246A  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/127,721  
 FILING DATE: 27-SEPTEMBER-1993  
 APPLICATION NUMBER: US 07/952,802  
 FILING DATE: 25-SEPTEMBER-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 6072035ak, Henry P.  
 REGISTRATION NUMBER: 33,200  
 REFERENCE/DOCKET NUMBER: 4-19276/A/P2/CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (908) 277-5110  
 TELEFAX: (908) 277-4306

## INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 424 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 22..402  
 FEATURE:

NAME/KEY: mat\_peptide  
 LOCATION: 82..402  
 OTHER INFORMATION: /product="light chain variable"  
 US-08-485-246A-7

Query Match  
 Best Local Similarity 47.4%; Score 45.4; DB 3; Length 424;  
 Matches 136; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 483 CAGGGGCCCCCTGCACCTTCAGCTTCGAGACCCCGAGCAGCAGCTGGAACTGAACCGGGA 542  
 DB 96 CCAGAGCCCCGGCAGCCCTGAGCCCTGAGCCCGGCGAGAGGGCCACCCCTGAGCTGCAGGGC 155  
 QY 543 GTGCGTGGCGCAGACCCCAACAGAACCATCTTGACAGTGAACCTGTATGATGTGCGAG 602  
 DB 156 CAGCCAGACATCGGCACCAACATCTGATGACAGCAGAACGCCGCGCAGGCCCCCGAG 215  
 QY 603 GCGGGGCGGCCCCCTGCCCCCAGCAGTAACTCGGGCATCAGGCGCAGCTGCTAGCGCAG 662  
 DB 216 GCTGCTGATCAAGTACGCGCAGCAGCAGCATCAGCGCATCCCGCAGAGGTTGAGCGCGAG 275  
 QY 663 GCGCGGGCGCATGAGGGGGCGCGCGCCTTACAGCAGGTCATCGGCGCATACCGCGG 722  
 DB 276 CGGCGAGCGGCGACGCTTCAACCTGACCATCAGCAGCTGAGAGCCCGAGAGACTTGGCCAT 335  
 QY 723 GTCCCTCTTCAGCAGCAGCAGCAGTGGGCGCGCCCTCTGCTGG 769  
 DB 336 GTACTACTGCCAGCAGCAGCAGCTGGGCCACCACTTCGGCCAGG 382

## RESULT 9

US-08-924-440-1  
 Sequence 1, Application US/08924440  
 Patent No. 5871550  
 GENERAL INFORMATION:

APPLICANT: Flits et al.  
 TITLE OF INVENTION: MUTANT THERMONOSPORA SPP. CELLULOSE  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genencor International, Inc.  
 STREET: 925 Page Mill Road  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSRO for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/924,440  
 FILING DATE: August 27, 1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:

FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Christopher L. Stone  
 REGISTRATION NUMBER: 35,696  
 REFERENCE/DOCKET NUMBER: GC388  
 TELECOMMUNICATION INFORMATION:



TELEPHONE: (650) 846-7555  
 TELEFAX: (650) 845-6504  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1293 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-924-440-1

Query Match 5.1%; Score 44; DB 2; Length 1293;  
 Best Local Similarity 47.5%; Pred. No. 0.3;  
 Matches 131; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 551 GCGCACCCCAAGACAGACATCTTCGACAGTGCATGATGATGTCGACAGGCGGCG 610  
 DB 263 GCGAACCCCAACGACATGACATCAACAGGCGCCCTGCGACAGACGCTCCGAGCGGGG 322  
 QY 611 GCGCCCTGCCCCCGACGATACCTGCGGCGATCAGCGCCGACGCTACGCGAGCGGCG 670  
 DB 323 GCGCCGCGGCTCCGGAACCCCTTCCTCCGACCCGCGACGACGCCCGGACCGGCGACCC 382  
 QY 671 GCATGAGAGGGGCGCGCCGACCTACAGAGGATGATGCGCCACTACCGGGGCTCTCT 730  
 DB 383 CGGTGAGCGGTACGCAAAAGTCCAGGTCTGCGGATCAGCTTGCGACGAGCAAGGCA 442  
 QY 731 TCCAGACACGACAGACAGTGGGCGCCCTCTCTGCTGGAGGAGACCGGCTCCACACA 790  
 DB 443 ACCCGGTCCAACTGCGCGGATAGACACCCAGCGGATCCAGTGGTCCGACCACTGCTGA 502  
 QY 791 CACACATCGCGCCCTAGAGAGCGAGCATCTGGA 826  
 DB 503 CCGACAGCTCGCTGAGACGCGCTGCGCTACGACTGGA 538

RESULT 10  
 US-08-804-227C-7  
 Sequence 7, Application US/08804227C  
 Patent No. 5876991

GENERAL INFORMATION:  
 APPLICANT: DeHoff, Bradley S.  
 APPLICANT: Kustross, Stuart A.  
 APPLICANT: Rostock, Paul R., Jr.  
 APPLICANT: Sutton, Kimberly L.  
 TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: THOMAS G. PLANT 1501  
 STREET: LILLY CORPORATE CENTER  
 CITY: INDIANAPOLIS  
 STATE: IN

COUNTRY: USA  
 ZIP: 46285  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: ASCII(DOS) Text only  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/804.227C  
 FILING DATE: February 21, 1997

CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Plant, Thomas, G.  
 REGISTRATION NUMBER: 35,784  
 REFERENCE/DOCKET NUMBER: X-8231  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 317-276-2459  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 44377 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single

TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 NAME/KEY: CDS  
 LOCATION: 350..14002  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 14046..20036  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 20110..31284  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 31329..36071  
 NAME/KEY: CDS  
 LOCATION: 36155..41830  
 US-08-804-227C-7

Query Match 5.1%; Score 44; DB 2; Length 44377;  
 Best Local Similarity 43.5%; Pred. No. 0.65; Mismatches 260; Indels 0; Gaps 0;  
 Matches 200; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 330 GCGCCGCGCTCGGCGCCACGACCGCGCTGCGCGCTTCGCGCCAGCGGAGCGCTT 389  
 DB 40615 GCGCGCGCGCGCGCGCGCGCGCGAGACACCCCGCGCGCGCTTCGCGCGCGCGCG 40674  
 QY 390 CCACGCGCTTCAGCCGACCTTATCGCTTACCTGCGACGAGATGACCTGCGACCACT 449  
 DB 40675 CGAGAGACCGCGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 40734  
 QY 450 CTGCGTGTGAGAGCGGAGGAGCGCCGACCCCTACAGAGGCGCGCTGACCGCTTGTG 509  
 DB 40735 CCAGTCCGAGAGCGCGCGCGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCG 40794  
 QY 510 GGACCCCGAGACGACGAGTGAACCGGAGGAGTGTGCGCGACCCCGCAAGAGAAC 569  
 DB 40795 GCTGCGCGCGGTGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 40854  
 QY 570 CATCTTCGACAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATG 629  
 DB 40855 CACCTTCGCTTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 40914  
 QY 630 TAAGTCGCGCATCAGCGGCGAGCGTGTAGGAGCGGCGCGCGCGCGCGCGCGCGCG 689  
 DB 40915 CGACGCGGAG 40974  
 QY 690 CACCTACAGAGAGTATGCGGCGCACTACCGGCGGCTTCCTTCAGACAGACAGAGAG 749  
 DB 40975 CTCCTTCG 41034  
 QY 750 TGGGCGCGCGCTTCCTGCTGAGAGGAGCGCGCTCCACAC 789  
 DB 41035 CCGTGGCGGCTTCGCGCGCGCGCTACTCCGCGAGCACCC 41074

RESULT 11  
 US-08-804-198-1  
 Sequence 1, Application US/08804198  
 Patent No. 5945320  
 GENERAL INFORMATION:  
 APPLICANT: Burgett, Stanley G.  
 APPLICANT: Kustross, Stuart A.  
 APPLICANT: Rao, Nagaraja R.  
 APPLICANT: Richardson, Mark A.  
 APPLICANT: Rostock, Paul R., Jr.  
 TITLE OF INVENTION: PLATENOLOIDE SYNTHASE GENE  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PAUL R. CANTRELL 1138  
 STREET: LILLY CORPORATE CENTER  
 CITY: INDIANAPOLIS  
 STATE: IN

COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh 7.0  
SOFTWARE: Microsoft Word 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,198  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CANTRELL, PAUL R.  
REGISTRATION NUMBER: 36,470  
REFERENCE/DOCKET NUMBER: P9113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3885  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4437 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 350..14002  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 14046..20036  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 20110..31284  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31329..36071  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 36155..41830  
US-08-804-198-1

Query Match 5.1%; Score 44; DB 2; Length 4437;  
Best Local Similarity 43.5%; Pred. No. 0.65;  
Matches 200; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 330 CGCCCGCCGCTGCGCCGACGCGCTGCGCGCCGCTTCCGCGGAGGCGCT 389  
DB 40615 CGCCCGCCGCTGCGCCGACGCGCGCGCGCTTCCGCGGAGGCGCT 40674  
QY 390 CCACGCGCTTCAGGCCACCTATCCGTACCTGACGACGAGATGACCTGCACACCAT 449  
DB 40675 CGAGGACGCGCGCGCGCTGCGCTGACCTGCTGCTCCGACAGTGGCGGCTCTCG 40734  
QY 450 CTCGCTGTGACAGGCGGAGAGAGCCCGACCCCTACGACGAGGCGCTTCAGCTTGG 509  
DB 40735 CCAGCTCGAGGAGCGCGCGGTGACGCGCGCGCGCTTCGCGGACCTGCGCTTCGACT 40794  
QY 510 GGACCCCGAGGACGACCTGGAATGAACGGGAGTGTGGCGGACCCCGCAAAAGAAC 569  
DB 40795 GCTCGCGCGGCTGCGCGCTGCGCGCTGCGCGGAGACACCGGCGCTGACCTGCGCG 40854  
QY 570 CATCTTCGACAGTGAATGATGATGACGAGGCTGGCGGCGCTTGCCTCCCGCCAGAG 629  
DB 40855 CACCGCTGCTTCGACGACGAGAGACCCCAACCGGCTGGCGGCGCAACACTGCGCGCTCG 40914  
QY 630 TAACTGGGCGATGAGCGCAAGTGTAGGAGGAGGCGGCGCGCATGAGAGGCGCGCGCC 689  
DB 40915 CGACGCGGAGGAGCGCGCGCGCGCGCGCGCGCTCGGCGCGAGAGGCGCTGTTCCG 40974  
QY 690 CACCTTACAGAGGTGATGAGGCGACACCGGCGCTCTCTTCGACACAGAGAGAGAG 749  
DB 40975 GTCTTTCGCGCGCGCTGGAACAGGAGGAGGCTGAGCGAGGCTGTGAGACTGATGGCGGA 41034

QY 750 TGGCGCCGCTCTGCTGAGGAGGAGCCGCGCTCCACAC 789  
DB 41035 CCGCGCGCGCTTCGCGCGCGCGCTACTCCGCGGAGACCC 41074

RESULT 12  
US-08-476-176B-5  
Sequence 5, Application US/08476176B  
Patent No. 5958708  
GENERAL INFORMATION:  
APPLICANT: Hardman, No. 5958708man  
APPLICANT: Kolbinger, Frank  
APPLICANT: Saldanha, Jose  
TITLE OF INVENTION: Reshaped monoclonal antibodies against an  
TITLE OF INVENTION: Immunoglobulin isotype  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5958708artis Patent Department  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07936-1080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,176B  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/127,721  
FILING DATE: 27-SEPTEMBER-1993  
APPLICATION NUMBER: US 07/952,802  
FILING DATE: 25-SEPTEMBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5958708ak, Henry P.  
REGISTRATION NUMBER: 33,200  
REFERENCE/DOCKET NUMBER: 4-19276/N/P2/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 277-5110  
TELEFAX: (908) 277-4306  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 22..402  
FEATURE:  
NAME/KEY: mat-peptide  
LOCATION: 82..402  
OTHER INFORMATION: /product="light chain variable  
OTHER INFORMATION: region C21-L1"  
US-08-476-176B-5

Query Match 5.1%; Score 43.8; DB 2; Length 424;  
Best Local Similarity 47.0%; Pred. No. 0.26;  
Matches 135; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 483 CCAGGCGCCGCTGACCTTCAGTTCGGGAGCCCGAGAGGAGCAAGTGAACCGGGA 542  
DB 96 CCAGAGCCCGCGGACCTGAGCTGAGCCCGCGGAGAGGCGCACCTGAGCTGAGCGGC 155  
QY 543 GTCCGTCGCGGACCCCGCAAGAGACATCTTCGAGAGTGAAGTGAATGATGAGTCCAG 602  
DB 156 CAGCGAGAGCATCGGACCAACATCATCTGTGACGAGAGAGCCCGGAGGCCCGCCAG 215



Query Match	5.1%;	Score 43.8;	DB 3;	Length 424;
Best Local Similarity	47.0%;	Pred. No. 0.26;		
Matches 135; Conservative	0;	Mismatches 152;	Indels 0;	Gaps 0;

```

? LENGTH: 424 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS

```

```

1 LOCATION: 22..402
2
3 FEATURE:
4 NAME/KEY: mat_peptide
5 LOCATION: 82..402
6 OTHER INFORMATION: /product="light chain variable
7 OTHER INFORMATION: region C21-L3"
8 US-08-127-721A-9
9
10 Query Match      5 1%; Score 43.8; DB 3; Length 424;
11 Best Local Similarity 47.0%; Pred. No. 0.26;
12 Matches 135; Conservative 0; Mismatches 152; Indels 0; Gaps
13
14 QY 483 CCAAGGCCCCCTGCACCCCTTCACACTTCGGGAGCCCGACAGCAGACTGGAACTGAACCGGGA 542
15      ||| ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
16 Db 96 CCAAGAGCCCGGACACCTTGAGCTTGAGGCCCGGAGAGGGCCACCCCTGAGCTGAGAGGC 155
17
18 QY 543 GTGCGTGCAGGCGACCCCAACAGAACCATCTTTCGACAGTGCATCTGATGATAGTGCCAG 602
19      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
20 Db 156 CAGCGAGAGATCGGAGCCACCATTCATCTGTGTACAGAGAACCCCGGACAGGCCCCAG 215
21
22 QY 603 GCTGGGCGGCCCTGCACCCCGCCAGAGAGTAATCGGGGATGAGGCCAGGTGTACGGAG 662
23      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
24 Db 216 GCTGCTGATCAGTACGTACGCCAGCAGAGCAATCAGGCGCATCCCGACAGGTTTGAAGGGAG 275
25
26 QY 663 CGGCGGCGCATGAGAGGGGCGCGCCCTACCTACAGAGAGTTCATCGGCACTAACCGGG 722
27      |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
28 Db 276 CGGCGACGGCAGCGACTTCACCTCGACATACAGAGGCTGGAGGCCGAGAGCTTCGCCAT 335
29
30 QY 723 GTCTCTCTTCAGACAGCAGAGCAGAGCAGTGGGCGCGCCCTCTTGCTGG 769
31      ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
32 Db 336 GTACTACTGCGACGAGAGCAGCACTTGCGCCACACCTTCGGGCCAG 382

```



Query Match	100.0%;	Score 861;	DB 10;	Length 1321,
Best Local Similarity	100.0%;	Pred. No. 1e-185;		
Matches 861; Conservative	0.	Microsoft		

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Sequence 55, Application US/09796753  
Publication No. US20030027998A1  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
FILE REFERENCE: 7853-227-999  
CURRENT APPLICATION NUMBER: US/09/796,753  
CURRENT FILING DATE: 2001-03-01  
PRIORITY APPLICATION NUMBER: 09/183,175  
PRIORITY FILING DATE: 1998-10-30  
PRIORITY APPLICATION NUMBER: 09/223,094  
PRIORITY FILING DATE: 1998-12-30  
PRIORITY APPLICATION NUMBER: 09/223,546  
PRIORITY FILING DATE: 1998-12-30  
PRIORITY APPLICATION NUMBER: 09/224,246  
PRIORITY FILING DATE: 1998-12-30  
PRIORITY APPLICATION NUMBER: 09/259,388  
PRIORITY FILING DATE: 1999-02-26  
PRIORITY APPLICATION NUMBER: 60/122,458  
PRIORITY FILING DATE: 1999-03-01  
PRIORITY APPLICATION NUMBER: 09/312,359  
PRIORITY FILING DATE: 1999-05-14  
PRIORITY APPLICATION NUMBER: 09/336,536  
PRIORITY FILING DATE: 1999-06-18  
PRIORITY APPLICATION NUMBER: 09/342,687  
PRIORITY FILING DATE: 1999-06-29  
PRIORITY APPLICATION NUMBER: 09/345,464  
PRIORITY FILING DATE: 1999-06-30  
PRIORITY APPLICATION NUMBER: 09/365,164  
PRIORITY FILING DATE: 1999-07-30  
PRIORITY APPLICATION NUMBER: 09/399,723  
PRIORITY FILING DATE: 1999-09-20  
PRIORITY APPLICATION NUMBER: 09/409,634  
PRIORITY FILING DATE: 1999-09-30  
PRIORITY APPLICATION NUMBER: 09/471,179  
PRIORITY FILING DATE: 1999-12-23  
PRIORITY APPLICATION NUMBER: 09/474,071  
PRIORITY FILING DATE: 1999-12-29  
PRIORITY APPLICATION NUMBER: 09/474,072  
PRIORITY FILING DATE: 1999-12-29  
PRIORITY APPLICATION NUMBER: 09/514,010  
PRIORITY FILING DATE: 2000-02-25  
PRIORITY APPLICATION NUMBER: 09/516,745  
PRIORITY FILING DATE: 2000-03-01  
PRIORITY APPLICATION NUMBER: 09/572,002  
PRIORITY FILING DATE: 2000-05-14  
PRIORITY APPLICATION NUMBER: 09/597,993  
PRIORITY FILING DATE: 2000-06-19  
PRIORITY APPLICATION NUMBER: 09/599,596  
PRIORITY FILING DATE: 2000-06-22  
PRIORITY APPLICATION NUMBER: 09/630,334  
PRIORITY FILING DATE: 2000-07-31  
PRIORITY APPLICATION NUMBER: 09/606,565  
PRIORITY FILING DATE: 2000-06-29  
PRIORITY APPLICATION NUMBER: 09/606,317  
PRIORITY FILING DATE: 2000-06-29  
PRIORITY APPLICATION NUMBER: 09/665,666  
PRIORITY FILING DATE: 2000-09-30  
PRIORITY APPLICATION NUMBER: 09/677,751  
PRIORITY FILING DATE: 2000-09-30  
NUMBER OF SEQ ID NOS: 162  
SEQ ID NO 55  
LENGTH: 969  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (6)...(761)  
US-09-796-753-55

Query Match 87.7%; Score 755.2; DB 9; Length 969;  
Best Local Similarity 99.6%; Pred. No. 7.7e-162;  
Matches 757; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 102 GGAGATCAGGAGCTGAGTTTGTTCATATCATCATCGTGGTGTGATGATGTTGAT 161  
Db 2 GGAGATGGGGGAGCTGGAGTTTGTTCATATCATCATCGTGGTGTGATGATGTTGAT 61  
QY 162 GGTGTGTGTGATCAGCTGCTGCTGAGCCACACAGAGTGTGTGACGGTCTTCATGAG 221  
Db 62 GGTGTGTGTGATCAGCTGCTGCTGAGCCACACAGAGTGTGTGACGGTCTTCATGAG 121  
QY 222 CCGGACAGCCAGGGGGGAGAGAGAAAGTCCCTGTCTCAGAGATGCTGTGACC 281  
Db 122 CCGGACAGCCAGGGGGGAGAGAGAAAGTCCCTGTCTCAGAGATGCTGTGACC 181  
QY 282 CTGGAGAGACAGTGTGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 341  
Db 182 CTGGAGAGACAGTGTGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 241  
QY 342 GCCCACCAGCCGCTGGGCGTGGCCCTTCCGCCAGGGGAGAGCCGCTTCCACCGTTCCA 401  
Db 242 GCCCACCAGCCGCTGGGCGTGGCCCTTCCGCCAGGGGAGAGCCGCTTCCACCGTTCCA 301  
QY 402 GCCCACCATATCCGTACTGCTGACAGCAGAGATGACCTGCCACCATCTGTGTCAGA 461  
Db 302 GCCCACCATATCCGTACTGCTGACAGCAGAGATGACCTGCCACCATCTGTGTCAGA 361  
QY 462 CGGGAG 521  
Db 362 CGGGAG 421  
QY 522 GCAGCTGAGTGAACGAGCCGAGATCGGTGCGGCGGACCCCAACAGAACATCTTGCAGAG 581  
Db 422 GCAGCTGAGTGAACGAGCCGAGATCGGTGCGGCGGACCCCAACAGAACATCTTGCAGAG 481  
QY 582 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641  
Db 482 TGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
QY 642 CAGCGCCAGCTGTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 701  
Db 542 CAGCGCCAGCTGTACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601  
QY 702 GGTATCGGCGGCTACCGCGGGGTCTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAG 761  
Db 602 GGTATCGGCGGCTACCGCGGGGTCTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAG 661  
QY 762 CTTCCTGAG 821  
Db 662 CTTCCTGAG 721  
QY 822 CTGGAGCAAGAGAGAGATTAACAGAAAGAGACACCTCTTC 861  
Db 722 CTGGAGCAAGAGAGAGATTAACAGAAAGAGACACCTCTTC 761

RESULT 4  
US-10-098-841-71  
Sequence 71, Application US/10099841  
Publication No. US20020197679A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Xu, Chongjun  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yunqing  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhao, Qing A.  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhilwei  
APPLICANT: Wehrman, Tom  
APPLICANT: Zhang, Jie

	Query Match	86.2%	Score 742.4;	DB 9	Length 1066;
	Best local Similarity	99.2%;	Pred. No.	6.1e-159;	Mismatches 746; Conservative 0; Mismatches 6; Indels 0; Gaps 0
OY	110 CGAGACGTGATCTTTCMTTCMAAATTCCATTTAAATCCTGGCAATTAATCGATTTTGCTTGCATTTTGAATGCTT				

	U	Indels	Gaps
QY	110	CGAGAGTGAAGTTTGTTCAGATCATCATCATTCGTGTGTGTATGATGATGGTATGGTATG	169
Db	113	CGAGACGACGATCTTGTTCAGATCATCATCATTCGTGTGTGTATGATGATGGTATGGTATG	169
QY	170	TGATCAGCGTCCGCGCGAGAGCCACTTCAAGCTGTGTGCACGGTCTCTTCATGAGCGGCACA	229
Db	173	TGATCAGCGTCCGCGCGAGAGCCACTTCAAGCTGTGTGCACGGTCTCTTCATGAGCGGCACA	229
QY	230	GCGAGGGGGGGAGAGAGAGATGCCCTGTCTCAGAAAGGATGCCCTGTGCCCCCTGGAGA	289
Db	233	GCGAGGGGGGGAGAGAGAGATGCCCTGTCTCAGAAAGGATGCCCTGTGCCCCCTGGAGA	289
QY	290	GCACAGTGTCAAGCGAAACGAATCCCAAGCGCGAGGTCTTACGCCGCCCTCGGCCACCG	349
Db	293	GCACAGTGTCAAGCGAAACGAATCCCAAGCGCGAGGTCTTACGCCGCCCTCGGCCACCG	349
QY	350	ACCGGCTTGGCCGTCGCCCTTGGCCAGCGGAGGGCTTCCAGCGGCTTCCAGGCCACCT	409
Db	353	ACCGGCTTGGCCGTCGCCCTTGGCCAGCGGAGGGCTTCCAGCGGCTTCCAGGCCACCT	409
QY	410	ATCCGTCACCTGCAGCAGAGATGACCTTGCACCCCAACATCTCGCTGTCAAGCGGGAG	469
Db	413	ATCCGTCACCTGCAGCAGAGATGACCTTGCACCCCAACATCTCGCTGTCAAGCGGGAG	469
QY	470	AGCCCCACCCCTTCAAGAGGCCCCCTGACCCCTTCAGACTTGGGAGCCCGACAGACAGCTGG	529
Db	473	AGCCCCACCCCTTCAAGAGGCCCCCTGACCCCTTCAAGACTTGGGAGCCCGACAGACAGCTGG	529
QY	530	AACCTGAACCGGAGGTGGGTGCGCGACCCCAACAGAGACATCTTGACAGTGA	589
Db	533	AACCTGAACCGGAGGTGGGTGCGCGACCCCAACAGAGACATCTTGACAGTGA	589
QY	590	TGGATAGTCCACAGGTGGGGGCGCCCTGCCCCCAGCAGTAATCTGGGCTATCAGCGCA	649
Db	593	TGGATAGTCCACAGGTGGGGGCGCCCTGCCCCCAGCAGTAATCTGGGCTATCAGCGCA	649
QY	650	CGTGTACGAGGAGCGGCGGCGATGAGAGGGCGCGGCCCACTACAGCGAGGTCAATCG	709
Db	653	CGTGTACGAGGAGCGGCGGCGATGAGAGGGCGCGGCCCACTACAGCGAGGTCAATCG	709
QY	710	GCCACTACCCGGAGTCTCTCTTCAGACACAGAGAGCAGTGTGGCGCCCTCTCTTCTGG	769
Db	713	GCCACTACCCGGAGTCTCTCTTCAGACACAGAGAGCAGTGTGGCGCCCTCTCTTCTGG	769
QY	770	AGGGAGCCCGGCTCCACACACACACATCGCGCCCTTGAAGAGGCGACGCACTTGGAGCA	829

Db 773 AGGGGACCCGGCTCCACACACACATCGGGCCCTTAGAGACCGAGCCATCTGGACA 833  
 Qy 830 AAGGAAGGATTAACAGAAGAAGACACCCCTTC 861  
 Db 833 AAGGAAGGATTAACAGAAGAAGACACCCCTTC 864

RESULT 5  
US-09-934-249-12

Patent No. US20020115081A1  
GENERAL INFORMATION

APPLICANT: Landschulz, Katherine T

APPLICANT: Thompson, John F.  
APPLICANT: Kennedy, Scott D

FILE REFERENCE: D0736 / 2001

CURRENT FILING DATE: 2001-08-21

PRIOR FILING DATE: 2000-08-22  
NUMBER OF SEQ ID NOS: 17

SEQ ID NO 12  
LENGTH: 878

ORGANISM: Mus musculus  
FEATURE:

LOCATION: (20) ... (841)  
S-09-934-248-13

Every Match	71.1%;	Score 612.4;	DB 10;	Length 878;
Best Local Similarity	86.5%;	Pred. No. 1.4e-129;		
Matches 701; Conservative	0.14e-129;			

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523</
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------



```

RESULT 6
US-10-000-256A-32
; Sequence 32, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-000-256A-32

```

Db	1125	TGCCCCCAGCACTAATCTGGGGCATTCAGGGCCAGCTGTTAGGGCAGGGGGGGCGCATG	1184
QY	676	GAGGGGCGCGCCGCCACCTCAAGCGAGGTATCGGCCCATACCGGGGTCTCTTCCAG	735
Db	1185	GAGGGGGCGCGGCCACCTCAAGCGAGGTATCGGCCCATACCGGGGTCTCTTCCAG	1244
QY	736	CACGAGCGAGCATTGGGCGCCCTCTTCTGGAGGGGACCGGGCTCCACACACACAC	795
Db	1245	CACGAGCGAGCATTGGGCGCCCTCTTCTGGAGGGGACCGGGCTCCACACACACAC	1304
QY	796	ATCGCGCCCTAGAGAGCGGACCATCTGGAGCAAAAGAGAGATAAACAGAAAGAAC	855
Db	1305	ATCGCGCCCTAGAGAGCGGACCATCTGGAGCAAAAGAGAGATAAACAGAAAGAAC	1364
QY	856	CCTCTC 861	
Db	1365	CCTCTC 1370	

```

1      RESULT 7
2      US-09-796-753-57
3      ; Sequence 57, Application US/09796753
4      ; Publication No. US20030027998A1
5      ; GENERAL INFORMATION:
6      ; APPLICANT: McCarthy, Sean A.
7      ; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
8      ; FILE REFERENCE: 7833-227-999
9      ; CURRENT APPLICATION NUMBER: US/09/796,753
10     ; CURRENT FILING DATE: 2001-03-01
11     ; PRIOR APPLICATION NUMBER: 09/183,175
12     ; PRIOR FILING DATE: 1998-10-30
13     ; PRIOR APPLICATION NUMBER: 09/223,094
14     ; PRIOR FILING DATE: 1998-12-30
15     ; PRIOR APPLICATION NUMBER: 09/223,546
16     ; PRIOR FILING DATE: 1998-12-30
17     ; PRIOR APPLICATION NUMBER: 09/224,246
18     ; PRIOR FILING DATE: 1998-12-30
19     ; PRIOR APPLICATION NUMBER: 09/259,388
20     ; PRIOR FILING DATE: 1999-02-26
21     ; PRIOR APPLICATION NUMBER: 60/1122,458
22     ; PRIOR FILING DATE: 1999-03-01
23     ; PRIOR APPLICATION NUMBER: 09/312,359
24     ; PRIOR FILING DATE: 1999-05-14
25     ; PRIOR APPLICATION NUMBER: 09/336,536
26     ; PRIOR FILING DATE: 1999-06-18
27     ; PRIOR APPLICATION NUMBER: 09/342,687
28     ; PRIOR FILING DATE: 1999-06-29
29     ; PRIOR APPLICATION NUMBER: 09/345,464
30     ; PRIOR FILING DATE: 1999-06-30
31     ; PRIOR APPLICATION NUMBER: 09/365,164
32     ; PRIOR FILING DATE: 1999-07-30
33     ; PRIOR APPLICATION NUMBER: 09/399,723
34     ; PRIOR FILING DATE: 1999-09-20
35     ; PRIOR APPLICATION NUMBER: 09/409,634
36     ; PRIOR FILING DATE: 1999-09-30
37     ; PRIOR APPLICATION NUMBER: 09/471,179
38     ; PRIOR FILING DATE: 1999-12-23
39     ; PRIOR APPLICATION NUMBER: 09/474,071
40     ; PRIOR FILING DATE: 1999-12-29
41     ; PRIOR APPLICATION NUMBER: 09/514,010
42     ; PRIOR FILING DATE: 2000-02-25
43     ; PRIOR APPLICATION NUMBER: 09/516,745
44     ; PRIOR FILING DATE: 2000-03-01
45     ; PRIOR APPLICATION NUMBER: 09/572,002
46     ; PRIOR FILING DATE: 2000-05-14
47     ; PRIOR APPLICATION NUMBER: 09/597,993
48     ; PRIOR FILING DATE: 2000-06-19
49     ; PRIOR APPLICATION NUMBER: 09/599,596
50     ; PRIOR FILING DATE: 2000-06-22
51     ; PRIOR APPLICATION NUMBER: 09/630,334
52     ; PRIOR FILING DATE: 2000-07-31
53

```

;; PRIOR APPLICATION NUMBER: 09/606,565  
;; PRIOR FILING DATE: 2000-06-29  
;; PRIOR APPLICATION NUMBER: 09/606,317  
;; PRIOR FILING DATE: 2000-06-29  
;; PRIOR APPLICATION NUMBER: 09/665,666  
;; PRIOR FILING DATE: 2000-09-20  
;; PRIOR APPLICATION NUMBER: 09/677,751  
;; PRIOR FILING DATE: 2000-09-30  
;; NUMBER OF SEQ ID NOS: 162  
;; SEQ ID NO 57  
;; LENGTH: 1713  
;; TYPE: DNA  
;; ORGANISM: Mouse  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (2)...(652)  
US-09-796-753-57

Query Match 53.7%; Score 462.2; DB 9; Length 1713;  
Best Local Similarity 83.8%; Pred. No. 1.4e-95;  
Matches 569; Conservative 0; Mismatches 83; Indels 27; Gaps 3;

QY 183 GCTGAGCCACTACAAAGCTGTCTGACGCTCTTCATCATGCCGGACACGCCAGGGGGAG 242  
DB 1 GCTGAGCCACTACAAAGCTGTCTGACGCTCTTCATCATGCCGGACACGCCAGGGGGAG 60  
QY 243 GAGAGAAAGTCCCTGTCTCTGAGAAAGATGCTGTGGCCCTCGGAGAGCAAGTGTGAG 302  
DB 61 GAGAGAAAGTCCCTGTCTCTGAGAAAGATGCTGTGGCCCTCGGAGAGCAAGTGTGAG 120  
QY 303 CAAGGAAATCCCAAGCGGAGGCTCTACGCGCCGCTGCGGCCACGAGCCGCTGCGCGT 362  
DB 121 ---TGGATGCGGAGGAGGAGGCTCTACGCGCCGCTGCGGCCACGAGCCGCTGCGCGT 177  
QY 363 GCGGCGCTTGGCCGAGGCGGAGGAGGCTCTACGCGCCGCTGCGGCCACGAGCCGCTGCG 422  
DB 178 GCGGCGCTTGGCCGAGGCGGAGGAGGCTCTACGCGCCGCTGCGGCCACGAGCCGCTGCG 228  
QY 423 GCAGGAATGAGCTGTCCACCCACACATCTGCTGTCTGAGAGGGGAGAGGCCACCCCTA 482  
DB 229 GCAGGAATGAGCTGTCCACCCACACATCTGCTGTCTGAGAGGGGAGAGGCCACCCCTA 288  
QY 483 CCAAGGCGCTTGGACCTTCAGCTGTGGGAGCCGCGAGCAAGTGTGAACTGAAACGGGA 542  
DB 289 CCAAGGCGCTTGGACCTTCAGCTGTGGGAGCCGCGAGCAAGTGTGAACTGAAACGGGA 348  
QY 543 GTGCGTCCGCGCAACCCCAAAACAGAACATCTTGCAGAGTGCATGATGATGATGCCAG 602  
DB 349 ATCTGTGCGCGCAACCCCTTAACCGAGCATCTTGCAGAGTGCATGATGATGATGATGCC 408  
QY 603 GTGCGGCGGCGCCGCGCGCCGCGAGCAATCTGCGGATCTGAGCGGCAAGTGTGAGCGG 662  
DB 409 GTGCGGCGGCGCCGCGCGCCGCGAGCAATCTGCGGATCTGAGCGGCAAGTGTGAGCGG 468  
QY 663 CCGCGGCGGCGATGAGAGGGGCGCGCGCCGCGAGCAATCTGAGCGAGTGTGAGCGGCA 722  
DB 469 CCGTGGGCGGATGAGAGGGGCGCGCGCCGCGAGCAATCTGAGCGAGTGTGAGCGGCA 528  
QY 723 GTCTCTCTTCCAGCAAGCAAGCAAGTGTGAGCGGCGCGCTCTGCTGTGAGAGGGAGCCG 782  
DB 529 GTCTCTCTTCCAGCAAGCAAGCAAGTGTGAGCGGCGCGCTCTGCTGTGAGAGGGAGCCG 588  
QY 783 CCACGACACACATGCGCGCCCTTGAAGAGCGAGCATCTGTGAGCAAGCAAGATGA 842  
DB 589 CCATGACTGTGCAATGTCGCCCACTGGA-----GAACAAAGGAGAGAGGAA 633  
QY 843 ACAGAAAGGACACCTCTC 861  
DB 634 ACAGAAAGGTCACCCCTC 652

RESULT 8  
US-09-934-249-14/c

;; Sequence 14, Application US/09934249  
;; Patent No. US20020115081A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lee, Richard T.  
;; APPLICANT: Landshultz, Katherine T.  
;; APPLICANT: Turtl, Thomas G.  
;; APPLICANT: Thompson, John F.  
;; APPLICANT: Kennedy, Scott P.  
;; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
;; FILE REFERENCE: P0738/7001/EP/KA  
;; CURRENT APPLICATION NUMBER: US/09/934,249  
;; PRIOR FILING DATE: 2001-08-21  
;; PRIOR APPLICATION NUMBER: US 60/227,159  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 14  
;; LENGTH: 693  
;; TYPE: DNA  
;; ORGANISM: Homo Sapiens  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (639)...(639)  
;; OTHER INFORMATION: a, c, g, or t/u  
US-09-934-249-14

Query Match 51.8%; Score 445.6; DB 10; Length 693;  
Best Local Similarity 97.2%; Pred. No. 6.3e-92;  
Matches 485; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 364 CCGCCCTTGGCCGAGGAGGAGGCTTCCACGCGCTTCCAGCCGCTATTCCTGACG 423  
DB 693 CCGCCCTTGGCCGAGGAGGAGGCTTCCACGCGCTTCCAGCCGCTATTCCTGACG 634  
QY 424 CACGAGATGACCTGTCCACCCACATCTGCTGTCTGAGAGGGGAGAGCCGCGTAC 483  
DB 633 CACGAGATGACCTGTCCACCCACATCTGCTGTCTGAGAGGGG--AGAGCCGCGTAC 577  
QY 484 CAGGCGCCCTGTACCTCTGAGCTTGGGAGCCCGAGCAGCAGCTGGAATGAAACGGGAG 543  
DB 576 CAGGCGCCCTGTACCTCTGAGCTTGGGAGCCCGAGCAGCAGCTGGAATGAAACGGGAG 517  
QY 544 TCGGTGCGCGCACCCCAAAACAGAACATCTTGCAGAGTGCATGAGTGAATGAGCCAG 603  
DB 516 TCGGTGCGCGCACCCCAAAACAGAACATCTTGCAGAGTGCATGAGTGAATGAGCCAG 457  
QY 604 CTGGCGGCGCCCTGCGCCCGCGAGCAATCTGCGGATCTGAGCGGCAAGTGTGAGCGG 663  
DB 456 CTGGCGGCGCCCTGCGCCCGCGAGCAATCTGCGGATCTGAGCGGCAAGTGTGAGCGG 397  
QY 664 GCGGCGGCGATGAGAGGGGCGCGCCGCGAGCAATCTGAGCGAGTGTGAGCGGCA 723  
DB 396 GCGGCGGCGATGAGAGGGGCGCGCCGCGAGCAATCTGAGCGAGTGTGAGCGGCA 337  
QY 724 TCCTCTTCCAGCAAGCAAGCAAGTGTGAGCGGCGCGCTCTGCTGTGAG--GGGAGCCG 782  
DB 336 TCCTCTTCCAGCAAGCAAGCAAGTGTGAGCGGCGCGCTCTGCTGTGAGGGGAGCCG 277  
QY 783 CCACGACACACATGCGCGCCCTTGAAGAGCGAGCATCTGTGAGCAAGCAAGATGA 842  
DB 276 CCGCCGACACACATGCGCGCCCTTGAAGAGCGAGCATCTGTGAGCAAGCAAGATGA 217  
QY 843 ACAGAAAGGACCTCTC 861  
DB 216 ACAGAAAGGACCTCTC 198

RESULT 9  
US-09-934-249-16  
;; Sequence 16, Application US/09934249  
;; Patent No. US20020115081A1  
;; GENERAL INFORMATION:



```

: TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
: FILE REFERENCE: PO-16.2c1
: CURRENT APPLICATION NUMBER: US/09/783,590
: CURRENT FILING DATE: 2000-02-15
: PRIOR APPLICATION NUMBER: 08/420,856
: PRIOR FILING DATE: 1995-04-12
: PRIOR APPLICATION NUMBER: 08/346,731
: PRIOR FILING DATE: 1994-11-21
: NUMBER OF SEQ ID NOS: 12485
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3488
: LENGTH: 65
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (51)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-3488

Query Match
Best Local Similarity 6.9%; Score 59.2; DB 10; Length 65;
Matches 61; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY 715 TACCCGGGGCTCTCTTCAGACGACGAGACGAGGGCCGCTTCCTGGAGGG 774
Db 1 TACCCGGGGCTCTCTTCAGACGACGAGACGAGAGCGGCGCCCTTCTGAGGGA 60
QY 775 ACCCG 779
Db 61 CCCGG 65

RESULT 13
US-10-184-644-332/c
: Sequence 332, Application US/10184644
: Publication No. US20030044930a1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jitan
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1C227
: CURRENT APPLICATION NUMBER: US/10/184,644
: CURRENT FILING DATE: 2002-06-28
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 332
: LENGTH: 520
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-184-644-332

Query Match
Best Local Similarity 5.3%; Score 46; DB 9; Length 520;
Matches 73; Conservative 79; Mismatches 210; Indels 0; Gaps 0;

QY 468 GGAGCCCCCACCACGACGAGGGCCCTGCACCCCTCAGCTTGAGGAGCCCCGAGCAGACT 527
Db 508 GHDDCWSMWAGS.AS.ACYG.BNH.W..CACTCBMR.TYCMSRC.KYG.B.THSN.W.H 449
QY 528 GGAACGTGAACCGGGAGTGTGTGCGCGACACCCCAACAGAACACATCTTCGACAGTGACCT 587
Db 448 HG.AACMCSSRR.TTCYCNSSCR.YBSRBN.CV.CM.CMRC.CSSCMRC.HCM.CT.C.T.389

```



Tue Mar 18 08:50:01 2003

us-09-934-249-3.rnpb

Page 10

OY 748 AGTGGGCCCTCTTGTGAGGAGGACCGGCTCCACCACACACATCGGCCCTTA 807  
Db 496 ACCCAGCGCATCCGCCCTGTGTGAGCACCCAGCTGTGTAACGCGACGCTGGCCGAG 555  
OY 808 GAGAGCGCAGCCATCTGGAGCAA 830  
Db 556 GAGGCGGTGTGTATCCGACCGCA 578

Search completed: March 17, 2003, 18:33:25  
Job time : 77.3941 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 11:34:54 ; Search time 1617.04 Seconds

(without alignments)  
8623.366 Million cell updates/sec

Title: US-09-934-249-3

Perfect score: 861

Sequence: 1 atgaccgcgttgatgggggt.....aacagaagacaccccttc 861

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estda:\*
- 2: em\_estdm:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hlc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hlc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_iny:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_ylt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_man:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	811.8	94.3	1007	9	AL578575
C 2	792.8	92.1	949	14	AL517150
C 3	743.6	86.4	967	14	BQ641849
C 4	700	81.3	1046	14	BM922276
C 5	678.6	78.8	805	9	AL558881
C 6	667.4	77.5	916	14	BQ954555

7	658.6	76.5	898	9	AL558882	AL558882
8	614.6	71.4	1079	11	BC023092	BC023092 Mus muscu
9	607.4	70.5	609	14	BQ636742	BQ636742 hml3n06.y
10	603.6	70.1	1207	11	AK008976	AK008976 Mus muscu
11	578.4	67.2	890	14	BQ690750	BQ690750 AGENCOURT
12	568.2	66.0	782	14	BQ015170	BQ015170 UT-H-ED1-
13	550	63.9	551	13	BM141979	BM141979 1f25a11.Y
14	529.8	61.5	894	13	BI851941	BI851941 603379004
15	519.4	60.3	729	14	BQ575741	BQ575741 UT-H-EZ1-
16	516.8	60.0	730	14	BM677602	BM677602 UT-H-E01-
17	501.6	58.3	588	13	BM483503	BM483503 536869.MA
18	468	54.4	857	12	BG323347	BG323347 602421734
19	461	53.5	763	13	BI646175	BI646175 603276395
20	445.6	51.8	693	9	AT761441	AT761441 w65f07.x
21	445.4	51.7	655	14	BQ681705	BQ681705 AGENCOURT
22	445.4	51.7	1280	14	BQ681500	BQ681500 AGENCOURT
23	438.6	50.9	825	9	AL543170	AL543170
24	430	49.9	841	13	BI156703	BI156703 602922119
25	428.8	49.8	974	10	BB624904	BB624904 BB624904
26	428.2	49.7	655	13	BI853324	BI853324 603379903
27	416.2	48.3	587	13	BI083462	BI083462 602875788
28	411	47.7	582	10	BE353323	BE353323 ut4h02.y
29	403	46.8	467	10	BE66930	BE66930 150581.MA
30	401.6	46.5	626	14	BM974296	BM974296 UT-E-EJ0-
31	400	46.5	626	14	BM712680	BM712680 UT-E-EJ0-
32	393	45.6	461	14	BM712680	BM712680 UT-E-EJ0-
33	392.2	45.6	990	12	BG675643	BG675643 602622053
34	392	45.5	464	12	BF026695	BF026695 601671765
35	389.8	45.3	633	14	BM714472	BM714472 UT-E-EJ0-
36	388	45.1	436	13	BM482193	BM482193 534988.MA
37	375.6	43.6	931	12	BF161499	BF161499 601770488
38	374	43.4	844	14	BQ686793	BQ686793 AGENCOURT
39	369.6	42.9	641	10	AM071693	AM071693 w53c08.x
40	369	42.5	559	12	BE855409	BE855409 7913f05.x
41	366.2	42.5	581	12	BG088400	BG088400 H3152F10-
42	365.6	42.5	577	12	BG075859	BG075859 H3152F10-
43	360.8	41.9	454	9	AA027926	AA027926 zk05h07.x
44	359.4	41.7	674	9	AI972096	AI972096 w62d11.x
45	358.8	41.7	570	14	BQ575582	BQ575582 UT-H-EZ1-

## ALIGNMENTS

RESULT 1  
AL578575/c 1007 bp mRNA linear EST 16-FEB-2001  
LOCUS AL578575 LTL\_NFL006.PL2 Homo sapiens cDNA clone CS0DK001YC24 3  
DEFINITION prime, mRNA sequence.  
ACCESSION AL578575  
VERSION AL578575.1 GI:12942781  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1007)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.  
location/Qualifiers  
1..1007  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DK001YC24"  
/clone\_id="LTL\_NFL006.PL2"  
/tissue\_type="placenta"  
/note="Vector: PCMVSPORT 6; Site\_1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end





QY 404 CCACCTATCCGCTACCTGACAGACAGATGACCTGACACCCACCATCTGCTGACAGC 463  
 |||||||  
 Db 590 CCACCTATCCGCTACCTGACAGACAGATGACCTGACACCCACCATCTGCTGACAGC 531  
 QY 464 GGGAG 523  
 |||||||  
 Db 530 GGGAG 471  
 QY 524 AGCTGGAATGAAACCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583  
 |||||||  
 Db 470 AGCTGGAATGAAACCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411  
 QY 584 ACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 643  
 |||||||  
 Db 410 ACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351  
 QY 644 GCGCAGCTGCTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 703  
 |||||||  
 Db 350 GCGCAGCTGCTACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 291  
 QY 704 TCATGCGGACCTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 763  
 |||||||  
 Db 290 TCATGCGGACCTACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 231  
 QY 764 TGCTGAG 823  
 |||||||  
 Db 230 TGCTGAG 171  
 QY 824 GGAGCAAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861  
 |||||||  
 Db 170 GGAGCAAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 133

RESULT 3  
 LOCUS B0641849 967 bp mRNA linear EST 15-JUL-2002  
 DEFINITION AGENCOURT\_8287174 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6292265  
 5', mRNA sequence.

ACCESSION B0641849  
 VERSION B0641849  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 967)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgsrps-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L1CM2493 row: 9 column: 18  
 High quality sequence stop: 571.  
 Location/Qualifiers

FEATURES  
 source  
 1..967  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6292265"  
 /clone\_lib="NIH\_MGC\_43"  
 /rname\_type="normal pigmented retinal epithelium"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: eye; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(c). Library constructed by Ling Hong

In the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH-MGC library.

BASE COUNT 194 a 334 c 296 g 143 t  
 ORIGIN

Query Match 86.4%; Score 743.6; DB 14; Length 967;  
 Best Local Similarity 99.5%; Pred. No. 1.1e-150;  
 Matches 746; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 112 GACCTGGAGATTGTTTCAGATCATCATCATGTTGTTGATGATGATGATGATGATG 171  
 Db 1 GACCTGGAGATTGTTTCAGATCATCATCATGTTGTTGATGATGATGATGATGATG 60  
 QY 172 ATACGTCCTGCTGACACCTACCAAGCTGTGACAGCTGCTTCATCAGCCGACAGC 221  
 Db 61 ATACGTCCTGCTGACACCTACCAAGCTGTGACAGCTGCTTCATCAGCCGACAGC 120  
 QY 232 CAGGGGGGAG 291  
 Db 121 CAGGGGGGAG 180  
 QY 292 ACAGTGTACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 351  
 Db 181 ACAGTGTACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 240  
 QY 352 CGCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411  
 Db 241 CGCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 412 CGGTACCTGACAGACAGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471  
 Db 301 CGGTACCTGACAGACAGAGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 472 CCCCCACCTTACAG 531  
 Db 361 CCCCCACCTTACAG 420  
 QY 532 CTGACAGCGGAG 591  
 Db 421 CTGACAGCGGAG 480  
 QY 592 GATAGTGCAGAGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 651  
 Db 481 GATAGTGCAGAGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
 QY 652 TGCTAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 711  
 Db 541 TGCTAGGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600  
 QY 712 CATACCCCGGGGCTCTCTTCCAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAG 771  
 Db 601 CATACCCCGGGGCTCTCTTCCAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 772 GGGACCGGGCTTCCAC 831  
 Db 661 GGGACCGGGCTTCCAC 720  
 QY 832 GAGAGAGATTAACAG 861  
 Db 721 GAGAGAGATTAACAG 750

RESULT 4  
 LOCUS BM922276 1046 bp mRNA linear EST 12-MAR-2002  
 DEFINITION AGENCOURT\_6707077 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5754437  
 5', mRNA sequence.  
 ACCESSION BM922276  
 VERSION BM922276  
 KEYWORDS EST.  
 SOURCE human.



D	b		805	GTCGCTGGTGCACCAACTGTCTTGACAGGTCCTTCAATCACCGGAACAGCCAGAGG	746
O	y		237	CGGAGAGAGAGAAGATGCCCTGTCTCAGAAAGATATGCTGTGGCCCTCGAGAGACAAGT	296
D	b		745	GGGGAGGAGAGAAAGATGGCTCTCTCTCAGAAAGATATGCTGTGGCCCTCGAGAGACAAGT	686
O	y		297	GTCCAGCAACGGAAATCCAGAGCCGAGTCTACGCCGCCCTCGGCCCATCCGACCCT	356
D	b		685	GTCCAGCAACGGAAATCCAGAGCCGAGTCTACGCCGCCCTCGGCCCATCCGACCCT	626
O	y		357	GGCCCTGGCCGCTTTCGCCCAGCGGGAGCGCTTCACCGCTTTCAGACCATTATCCGTA	416
D	b		625	GGCCCTGGCCGCTTTCGCCCAGCGGGAGCGCTTTCAGACCATTATCCGTA	566
O	y		417	CCTGCAGACAGAGATGACCTTCGACCCAGCAACCATCTCGCTGTGAGAGGGAGAGCCCC	476
D	b		565	CCTGCAGACAGAGATGACCTTCGACCCAGCAACCATCTCGCTGTGAGAGGGAGAGCCCC	506
O	y		477	ACCCATACAGAGGGCCCTGACACCTCAGTTCTGGAGACCCCGAGCGACAGCTGGAACTGAA	536
D	b		505	ACCCATACAGAGGGCCCTGACACCTCAGTTCTGGAGACCCCGAGCGACAGCTGGAACTGAA	446
O	y		537	CCGGGAGTCGTGTGCGCGGACCCCGCAAAGAACCATCTTCGACAGTAGCTATGATGATAG	596
D	b		445	CCGGGAGTCGTGTGCGCGGACCCCGCAAAGAACCATCTTCGACAGTAGCTATGATGATAG	386
O	y		597	TGCCAGGCTGTGGGGGCCCCCTGTGCCCGACAGTAACTCGGGCATCAGCGCCAGCTGCTA	656
D	b		385	TGCCAGGCTGTGGGGGCCCCCTGTGCCCGACAGTAACTCGGGCATCAGCGCCAGCTGCTA	326
O	y		657	CGGACGCGGGGGGCGCATATGAGAGGGGCGCGGCCACCTACAGCGAGGTATATGGCCACTA	716
D	b		325	CGGACGCGGGGGGCGCATATGAGAGGGGCGCGGCCACCTACAGCGAGGTATATGGCCACTA	266
O	y		717	CCCGGGGTCTCTCTTCCAGACACACAGACAGTAGTGGGCGCCCTCTTGCTGGAGGGAGC	776
D	b		265	CCCGGGGTCTCTCTTCCAGACACACAGACAGTAGTGGGCGCCCTCTTGCTGGAGGGAGC	206
O	y		777	CCGGCTCCACACACACACATCGCGGCGCCCTTAGAGAGCGGACCATCTGGAGCAAGAGAA	836
D	b		205	CCGGCTCCACACACACACATCGCGGCGCCCTTAGAGAGCGGACCATCTGGAGCAAGAGAA	146
O	y		837	GGATAACAGAAAGACACCCCTCTC	861
D	b		145	GGATAACAGAAAGACACCCCTCTC	121
S	E	S	LOCUS		
S	E	S	B0954555	916 bp mRNA linear EST 21-AUG-2002	
S	E	S	DEFINITION	AGNESCOURT_8825282 lupski.scfalic_nerve Homo sapiens CDNA clone IMAGE:6204609 5', mRNA sequence.	
S	E	S	ACCESSION	B0954555	
S	E	S	VERSION	B0954555.1 GI:22370033	
S	E	S	KEYWORDS	EST.	
S	E	S	SOURCE	human.	
S	E	S	ORGANISM	Homo sapiens	
S	E	S		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
S	E	S		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
S	E	S	REFERENCE	NIH-MGC http://mgc.ncl.nih.gov/.	
S	E	S	AUTHORS	1 (bases 1 to 916)	
S	E	S	TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
S	E	S	JOURNAL	Unpublished (1999)	
S	E	S	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@db-pm.nlm.nih.gov Tissue Procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> plate: LLNL1626 row: C column: 10	

	FEATURES	High quality sequence stop: 669.
	source	Location/Qualifiers
OY	/organism="Homo sapiens"	1..916
Dd	/db_xref="taxon:9606"	
OY	/clone="IMAGE:6204609"	
Dd	/clone_lib="lupski_sciatic_nerve"	
OY	/sex="male"	
Dd	/tissue_type="sciatic nerve"	
OY	/dev_stage="adult, 70 yr"	
Dd	/lab_host="DH10B"	
OY	/note="vector: pcMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; CDNA made by oligo-dT priming. Directionally cloned using the following adaptors:	
Dd	'-TCGACCCAGCGGTCTGC'-3' and 5'-ACTATGTTCTAGATCAGCAAGGGGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."	
BASE COUNT	162 a      354 c      273 g      127 t	
ORIGIN		
Query Match	77.5%; Score 667.4; DB 14;	Length 916;
Best Local Similarity	95.3%; Pred. No. 3.2e+14;	
Matches 731; Conservative	0; Mismatches 31; Indels	5; Gaps 4,
OY	1 ATGCACCGGTTATATGGGGTGCAAACAGCACCGCGCGCGCGCGGAGCCCATAATGTC 60	
Dd	70 ATEGACCCTTGATTATGGGGGTCAAACAGCACCGCGCGCGCGCGGAGCCCATAATGTC 129	
OY	61 TCCCTCAGGTGCAATGCAACAACGCTCTTGTGTCCAAGAGCATGAGATACGAGGCTGGAG 120	
Dd	130 TCCTCAGGTGCAATGCAACAACGCTCTTGTGTCCAAGAGCATGAGATACGAGGCTGGAG 189	
OY	121 TTGTGTTCAGATCATCATCTCGTGGTGTGATGATGCTGATGTGCTGATCAACGTTC 180	
Dd	190 TTGTGTTCAGATCATCATCTCGTGGTGTGATGATGATGATGTGCTGATCAACGTTC 249	
OY	181 CTGCTGAGCACTAACAGGTGTCTGACAGCGTCTTCATATAGCGCGGACAGCCAGAAGGGCGG 240	
Dd	250 CTCGTGAGCACTAACAGGTGTCTGACAGCGTCTTCATATAGCGCGGACAGCCAGAAGGGCGG 309	
OY	241 AGGAGAGAAGATGCCCCCTGTCTCAGAAGATGCTGTGGCCCTCGAGAGACAGTAGTGA 300	
Dd	310 AGGAGAGAAGATGCCCCCTGTCTCAGAAGATGCTGTGGCCCTCGAGAGACAGTAGTGA 369	
OY	301 GGCAACGGATCCACAGAGCCGAGGTTCTACGCCCGCCCTCGGCCACAGACCGCTGGCC 360	
Dd	370 GGCAACGGATCCACAGAGCCGAGGTTCTACGCCCGCCCTCGGCCACAGACCGCTGGCC 429	
OY	361 GTGCGCGCCCTTGCSCACAGGGGAGGGGCTTCACGCCCTTCACGCCCACTATCCGTACCTG 420	
Dd	430 GTGCGCGCCCTTGCSCACAGGGGAGGGGCTTCACGCCCTTCACGCCCACTATCCGTACCTG 489	
OY	421 CAGCAGAGATGACCTGCACACCAACATCTGCTGTGAGAGGGGAGAGGCCCCACCC 480	
Dd	490 CAGCAGAGATGACCTGCACACCAACATCTGCTGTGAGAGGGGAGAGAGGCCCCACCC 549	
OY	481 TAACAGGGCCCCCTGCACCTCTCAGCTTGCGGAGCCCGAGCAGCAGCTGGAATGAACGG 540	
Dd	550 TAACAGGGCCCCCTGCACCTCTCAGCTTGCGGAGCCCGAGCAGCAGCTGGAATGAACGG 609	
OY	541 GAGTGGGTGGCGGACCCCAACAGAACCACTTTCGACAGTAGTACTGATGTAGTGTGCG 600	
Dd	610 GAGTGGGTGGCGGACCCCAACAGAACCACTTTCGACAGTAGTACTGATGTAGTGTGCG 669	
OY	601 -AGGCTGGAGGGCCCCCCAGAGAGTAATCTGGGCAATCAG-CGCAACGTCTACTAG 658	
Dd	670 AAAGGCTGGAGGGCCCCCCAGAGAGTAATCTGGGCAATCAGCGCCACGTCTACTAG 729	
OY	659 GCAGCGGGGGGG-GCATGGAGGGGGCGCGCCGACCACTACAGCGAGGTATCGGCA--CT 715	

RESULT	7
AL558862	
LOCUS	
DEFINITION	898 bp mRNA
AL558862	LT1_NFL008_Tc2 Homo sapiens cDNA clone GSD0J015YF12 5
ACCESSION	prime, mRNA sequence.
AL558862	
AL558862.1	GI:12903858
KEYWORDS	EST.
SOURCE	human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eultheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 898)  
L14W.B., Gruber,C., Jassse,J., and Polajays,D.  
Full-length cDNA libraries and normalization  
unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
JOURNAL location/Qualifiers  
1. 898

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0J015F12"
/clone_lib="LTI_NFL008_TC2"
/sex="male"
/class="T"
/class_type="T"
/notes="Vector: pCMVSPORT 6; site_1: NotI; 1st strand
enriched with a NotI-oligo(dT) primer. Five prime
cloned into the Not I and Eco RV sites of the pCMVSPORT
life technologies. Contact : Feng Liang Life Technology
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email: fliang@lifestechnology.com URL :
http://fulllength.invitrogen.com"

```

	Query Match	76.5%;	Score 658.6;	DB 9;	Length 898;
	Best Local Similarity	99.3%;	Pred. No. 2.6e-132;		
	Matches 680;	Conservative	2;	Mismatches 1;	Indels 2;
					Gaps
QY	177	GTGCTGCTGACGCACATACAGCTGTGTGACAGGCTTCATGACGGGACACAGCCAGG	236		
Db	1	GTGCTGTGTGACGCACATACAGCTGTGTGACAGGCTTCATGACGGGACACAGCCAGG	236		
QY	237	GCGAGAGAGAAAGATGCTCCCTCTCTCGAAGATGCTGTGGCCCTCGAGAGACAGT	296		
Db	61	GCGAGAGAGAAAGATGCTCCCTCTCTCGAAGATGCTGTGGCCCTCGAGAGACAGT	296		
QY	297	GTCAGAGAACGAGATGCCAAGACCCGACAGGTCTAGGCCCGCTCGGGCCACGACGCGCT	356		
Db	121	GTCAGAGAACGAGATGCCAAGACCCGACAGGTCTAGGCCCGCTCGGGCCACGACGCGCT	356		
QY	357	GCGCGTGTCCGCCCTTGCCACGCGGGAGGCTCTCCACGCGCTTCCAGGCCACATTCGTA	416		
Db	180	GCGCGTGTCCGCCCTTGCCACGCGGGAGGCTCTCCACGCGCTTCCAGGCCACATTCGTA	416		
QY	417	CCGCGACGACGAGATGACCTGCCACCCACATCTGCTGTCAAGCGGGAGAGGCCCC	476		
Db	240	CCGCGACGACGAGATGACCTGCCACCCACATCTGCTGTCAAGCGGGAGAGGCCCC	476		

QY	4777	ACCCATACAGAGGCCCTCTGCAACCTTCAGGCTTGGGAGCCCCGAGCGAGCGAGTGGAACTTGA	536
Db	300	ACCTTACAGAGGCCCTCTGCAACCTTCAGGCTTGGGAGCCCCGAGCGAGCGAGTGGAACTTGA	536
QY	537	CCGGGAGTCCGATCGCGGCAACCCCAACGAAACATCTTCCAGCATGTAACCTGATGATAG	536
Db	360	CCGGGAGTCTMGATCGCGGCAACCCCAAGAAACATCTTCCAGCATGTAACCTGATGATAG	536
QY	597	TGCAGAGTCCGCGCGGCCCTTGCCTCCCGCCGACGTAACCTCGGGCATCGAGCGCAGCTGCTA	656
Db	420	TGCAGAGTCCGCGCGGCCCTTGCCTCCCGCCGACGTAACCTCGGGCATCGAGCGCAGCTGCTA	656
QY	657	CGGACATCGGCGGGGCGATGAGAGGGGCGCGCCACCTTACAGGAGGATATCGGCGACTA	716
Db	480	CGGACATCGGCGGGGCGATGAGAGGGGCGCGCCACCTTACAGGAGGATATCGGCGCGACTA	716
QY	717	CCCCGGGCTCTCTTCACGACACGACGAGAGATGGGGCGCCCTCTTCTGAGAGGGAGC	776
Db	540	CCCCGGGCTCTCTTCACGACACGACGAGAGAGATGGGGCGCCCTCTTCTGAGAGGGAGC	776
QY	777	CCGGCTCCACACACACACATCGCGGCCCTTACGAGAGCGAGGCATCTGGAGCAAAAGAGAA	836
Db	600	CCGGCTCCACACACACACATCGCGGCCCTTACGAGAGCGCA-SCATCTGGAGCAAAAGAGAA	836
QY	837	GGATAAACAGAAAGAGACCTTCTC	861
Db	659	GGATAAACAGAAAGAGACCTTCTC	683

RESULT	8
BC023092	
LOCUS	
DEFINITION	BC023092
ACCESSION	Mus musculus, Nedd4 WW binding protein 4, clone IMAGE:3989996,
VERSION	mRNA.
KEYWORDS	BC023092
SOURCE	BC023092.1 GI:18605637
ORGANISM	HPC. house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 1079)
AUTHORS	Strasberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (04-FEB-2002) NCBI

REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

Email: cgephs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Guarinate, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, C.A.  
  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Series: IRMK Pattern: 36 Row: h Column: 6  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Similarity but not  
 identity to protein  
 This clone has the following problem: frame shifted.  
 Location/Qualifiers  
 1..1079  
 /organism="Mus musculus"

```

/db_xref="taxon:10090"
/map="CECH II"
/clone="IMAGE:3989996"
/tissue_type="Mammary tumor metastatized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MMTV."
/clone_id="NCI_CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT      271 a      341 c      289 g      178 t
ORIGIN

```

```

Query Match      71.4%; Score 614.6; DB 11; Length 1079;
Best Local Similarity 86.8%; Pred. No. 8.6e-123;
Matches 702; Conservative 0; Mismatches 95; Indels 12; Gaps 2;

```

```

QY 5 ACCGCTGATGAGGGGTCAACAGCACCGCCGCCGCCGCGGAGCCCAATGTCCT 64
DB 15 ACCGCTGATGAGGGGTCAACAGCACCGCCGCCGCCGCGGAGCCCAATGTCCT 74
QY 65 GCAGGTGCAATGGAACGGCTTGTGTCAGAGCATGAGATGACGAGGAGTTTG 124
DB 75 GCGGTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 134
QY 125 TTCAGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 184
DB 135 TGCAGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 194
QY 185 TGAGCCACTACAGCTGTCTGACAGCTGTCTGACAGCTGTCTGACAGCTGT 244
DB 195 TGAGCCACTACAGCTGTCTGACAGCTGTCTGACAGCTGTCTGACAGCTGT 254
QY 245 GAGAGAGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304
DB 255 GAGAGAGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
QY 305 ACGGATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
DB 313 -TGGAAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 371
QY 365 CGCCCTTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
DB 372 CCCCCTTATCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 422
QY 425 ACGGATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
DB 423 ACGGATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
QY 485 ACGGATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
DB 483 ACGGATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
QY 545 CGGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
DB 543 CTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 602
QY 605 TGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
DB 603 TGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
QY 665 GCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
DB 663 GTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
QY 725 CTTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
DB 723 CTTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
QY 785 ACGACAGACAGATGCGGCGCTAGAGAG 813
DB 783 ATCACTGACACATGCGGCGCTAGAGAG 811

```

```

RESULT 9
B0636742
LOCUS
DEFINITION
hd13h06.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
ACCESSION
B0636742
VERSION
B0636742.1 GI:21761201
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 609)
Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman
J.W., Bouffard, G., Smith, D., and Peterson, K.
Expressed sequence tag analysis of human retina for the NEBank
Project: Retblindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
Mol. Vis. 8 (4), (2002) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 13 row: h column: 06
Seq primer: M13Rpl reverse primer (ABI).
Location/Qualifiers
1..609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hd13h06"
/clone_id="Human Retina cDNA (Un-normalized, unamplified)"
): hd/he
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the Superscript Plasmid System full
details of which are contained in the manufacturer's
instruction manual. (http://www.lifetech.com/). First
strand synthesis was carried out using a Not I
primer-adaptor (5'-TGACTAGTTCATGATCGGAGCGCGCCGCTT)15-3'
}. EST analysis was performed on the unamplified library
at the NIH Intramural Sequencing Center (NISC)."
BASE COUNT      114 a      238 c      182 g      75 t
ORIGIN
Query Match      70.5%; Score 607.4; DB 14; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.8e-121;
Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

AUTHORS	Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Iwata, A., Fukunishi, Y., Konno, H., Adachi, Y., Fukuda, S., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasuyama, T., Yamazaki, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiya, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Okazakunishi, J., Schiml, L. M., Stabill, F., Suzuki, R., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldreich, R., Bersh, G., Blake, J., Boffelli, D., Bologna, N., Cernici, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, T., Mombaerts, P., Nardone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyokawa, K., Wang, K. H., Weltz, C., Wiltaker, C., Williams, L., Wynshaw-Borris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 403 (6821), 685-690 (2001)
PMID	21085560
PUBMED	11217851
REFERENCE	5 (bases 1 to 1207)
AUTHORS	Adachi, J., Aizawa, K., Akita, Y.,

HARA, A., Fukuda, S., Fukunishi, Y., Grawestien, M., Bult, C., Hayatsu, N., Hill, D., Hiramoto, F., Furuno, M., Hasegaki, T., Hime, D., Imotochi, K., Ishii, Y., Itoh, M., Izawa, T., Horii, F., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Kusaka, T., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishii, K., Nomura, R., Ohno, M., Okazaki, I., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shimizu, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, D., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, T., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yamazaki, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (10-III-2000).

RIKEN Yokohama Research Group, RIKEN Genomic Laboratory for Genome  
RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,  
URI: <http://genome.gsc.riken.go.jp/>,  
Fax: 81-45-503-9216, Tel: 81-45-503-9222,  
COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared

Division of Experimental Animal Research Center and Genome Science Laboratory in RIKEN  
[3 GAGGACAGAGCATCCAGACCTCTTTTGTTCATTTTTCVN 3'] cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rct = 5.0 and tapper. cDNA went = 25.0 Second strand

```

GAGGAGAGACATTCGACGTTAATT7AAATTAATCCCCCCCCCCCCC 3'}. cDNA was cloned
with XhoI and SctI. Cloning sites, 5' end: XhoI; 3' end: SctI.
Host: SOLR.
Location/Qualifiers
1..1207
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2210418102"
/db_xref="MGI:MGI:1902457"
/db_xref="taxon:10090"

```

```

/tissue_type="stomach"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

```

```

CDS
      /dev_stage="adult"
      391..1173
      /note="Nedd4 WW binding protein 4
      data source:MGI, source key:MGI:1929600, evidence:ISS
      putative"
      /codon_start=1
      /protein_id="BAB26001.1"
      /db_xref="GI:12843489"
      /db_xref="MGI:1929600"
      /translation="MSPARATADRSLEPSEITELEPVQIVIVVMVMVWVITCL
      SHKLARSTISHSQARRRQDGLSSBGCWPSSESVSGMPPEPVYAPPRDRLAV
      PPIQSRREPTPYLQHEJALPPTISLSGSEPPPOGCTQLRDEQOLENRES
      VRAPNNRTFSDLDISTMLGSCPPSSNGISATCYSSGSRGMPPTYSVEVGHYP
      GSSFOHOOSNGPSLSLEGTRLHSHIAPLENKEKXOKKGPL"
BASE COUNT      200 a      459 c      374 g      173 t      1 others
ORIGIN
Query Match      70.1%; Score 603.6; DB 11; Length 1207;
Best Local Similarity 85.2%; Pred. No. 2.1e-120;
Matches 731; Conservative 0; Mismatches 99; Indels 28; Gaps 4;
QY 5 ACCGCTGATGGGGGTCAACAGCACCGCCGCCGCCGCCGCCGCAATGTCTCT 64
DB 340 ACCGCTGATGGGGGTCAACAGCACCGCCGCCGCCGCCGCCGCAATGTCTCT 399
QY 65 GCACGTGCACTG-CAAACGCTTTGTTCAGAGACATGGAATCAGGAGCTGAGTT 123
DB 400 GCGGTGCACTGCCCAAGCGCTCTTTGTTCAGAGACATGGAATCAGGAGCTG 459
QY 124 GTTCAGATCATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183
DB 460 GTTCAGATCATCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 519
QY 184 CTGAGCCATCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 243
DB 520 CTGAGCCATCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
QY 244 AGAAGATGCGCTGTCTCTGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGT 303
DB 580 AGAAGATGCGCTGTCTCTGAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGT 638
QY 304 AACGGAATCCAGAGCGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 363
DB 639 --TGAATGCGGAGCGACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 696
QY 364 CCGGCTTCCGCCAGCGGGAGCGCTTCCAGCGCTTCCAGCGCTTCCAGCGCTT 423
DB 697 CCGGCTTCCAGCGGGAGCGG-----AGCGGATTCCAGCGCTTCCAGCGCTT 747
QY 424 CAGAGATGCACTGCGACCCAGCATCTCGGTTCAGAGCGGAGAGCGGAGCGGAG 483
DB 748 CAGAGATGCGCTGCGACCCAGCATCTCACTGTCTGTGTGTGTGTGTGTGTGT 807
QY 484 CAGGCGCCCTGCACCTTCAGCTTCGGAGCCCGGAGCGAGCAAGCTGGAATGAA 543
DB 808 CAGGCGCCCTGCACCTTCAGCTTCGGAGCCCGGAGCGAGCAAGCTGGAATGAA 867
QY 544 TCGGTGGCGGACCGCCCAACAGCAATCTTTCAGAGTGAAGCTGATGATGCTCA 603
DB 868 TCGGTGGCGGACCGCCCAACAGCAATCTTTCAGAGTGAAGCTGATGATGCTCA 927
QY 604 CTGGGCGGCGCTTCCGCCAGGATGATGCGGAGCATGAGCGGAGCTGATGAGGAG 663
DB 928 CTGGGCGGCGCTTCCGCCAGGATGATGCGGAGCATGAGCGGAGCTGATGAGGAG 987
QY 664 GCGGCGGCGATGAGGAGCGCGCCCACTTACAGCAAGGTCACTGAGCAATCCCGGG 723
DB 988 GGTGGGCGCATGAGGAGCGCGCCCACTTACAGCAAGGTCACTGAGCAATCCCGGG 1047
QY 724 TCCCTCTTCCAGACAGAGAGAGTGGCGCGCTCTCTGTGTGTGTGTGTGTGTGT 783
DB 1048 TCCCTCTTCCAGACAGAGAGAGTGGCGCGCTCTCTGTGTGTGTGTGTGTGTGT 1107

```

```

QY 784 CACCACACACATCGCGCCCTGAGAGAGCGACGACCATCTGAGCAAAAGAGATAAA 843
DB 1108 CATCACTGCACTATTCGCCCACTGGA-----GAACMAAGAGAGAGAGAAA 1152
QY 844 CAGAAAGCACACCTCTC 861
DB 1153 CAGAAAGTCAACCCCTC 1170
RESULT 11
B0690750
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: egapds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LCM2368 row: h column: 14
High quality sequence stop: 627.
location/Qualifiers
1..890
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1db="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit.
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      168 a      303 c      273 g      146 t
ORIGIN
Query Match      67.2%; Score 578.4; DB 14; Length 890;
Best Local Similarity 93.5%; Pred. No. 5.0e-115;
Matches 615; Conservative 0; Mismatches 41; Indels 2; Gaps 1;
QY 87 TTTGTCAGAGATGAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGT 146
DB 105 TCTCTCGAAGAACAGGCAATGGCGAGCTGAGATTTGTTTCAGATCATCATCATG 164
QY 147 GGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 206
DB 165 GGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 224
QY 207 ACGGTCTTATATGCGCGGACAGCCAGCGGCGGAGAGAGAGAGATGCGCTGTCTG 266
DB 225 ACGGTCTTATATGCGCGGACAGCCAGCGGCGGAGAGAGAGATGCGCTGTCTG 284
QY 267 AGGATGCTGTGCGCTCGGAGAGACAGTGTCAAGCAAGGATCCAGAGCGGAGGT 326

```









RESULT 15  
 B0575741/c  
 LOCUS  
 DEFINITION  
 B0575741  
 UI-H-E21-dbg-h-14-0-UI.s1 mRNA linear EST 19-JUN-2007  
 UI-H-E21-dbg-h-14-0-UI.3', mRNA sequence.  
 ACCESSION  
 VERSION  
 B0575741  
 B0575741.1 GI:21479058  
 KEYWORDS  
 ESR.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (Bases 1 to 729)  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 TITLE  
 http://www.ncbi.nlm.nih.gov/cgicgap  
 JOURNAL  
 Tumor Gene Index  
 COMMENT  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov

Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of Orthopaedics  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD  
Polimers:  
Location/Qualifiers  
I. .729

```

BASE COUNT
ORIGIN
101 a
213 c 253 g 162 t

```

[illegible]

Tue Mär 18 08:50:02 2003

us-09-934-249-3.rst

Page 13

**Oy** 821 TCTGGAGCAAAAGAGAAGATTAACAGAAAAAGCACCCCTCTC 861  
| | | | | | | | | | | | | | | | | | | | | |  
**Db** 249 TCTGGAGCAAAAGAGAAGATTAACAGAAAAAGCACCCCTCTC 209

Search completed: March 17, 2003, 13:40:16  
Job time : 1621.04 secs

**THIS PAGE BLANK (USPTO)**

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result	Query	Result	Query
Success	SELECT * FROM ...	Success	SELECT * FROM ...
Failure	SELECT * FROM ...	Failure	SELECT * FROM ...

[illegible]

2	1229	93.0	4839	9	AF305616	AF305616	HC
3	0874	67.3	1061	0	PC015019	PC015019	WC

Case	Age	Sex	Site	Pathologic	Survival
4	861	65.2	861	6	AX392419
5	803	60.7	1141	8	AX392419
6	803	60.7	1141	8	AX392419

6	800	60.6	1818	9	AY128643	AY128643	HC
7	616.4	46.7	878	5	AY202438	AY202438	SC

Case	Age	Sex	Occupation	Education	Marital Status	Religion	Political Party	Income	Assets	Liabilities	Net Worth	Spouse's Net Worth	Children's Net Worth	Other Assets	Other Liabilities	Other Net Worth	Other Spouse's Net Worth	Other Children's Net Worth	Other Assets	Other Liabilities	Other Net Worth	Other Spouse's Net Worth	Other Children's Net Worth
8	594.4	45.0	61505	9	AF305426																		
9	594.4	45.0	61505	9	AF305426																		
10	594.4	45.0	61505	9	AF305426																		
11	594.4	45.0	61505	9	AF305426																		
12	594.4	45.0	61505	9	AF305426																		
13	594.4	45.0	61505	9	AF305426																		
14	594.4	45.0	61505	9	AF305426																		
15	594.4	45.0	61505	9	AF305426																		
16	594.4	45.0	61505	9	AF305426																		
17	594.4	45.0	61505	9	AF305426																		
18	594.4	45.0	61505	9	AF305426																		
19	594.4	45.0	61505	9	AF305426																		
20	594.4	45.0	61505	9	AF305426																		
21	594.4	45.0	61505	9	AF305426																		
22	594.4	45.0	61505	9	AF305426																		
23	594.4	45.0	61505	9	AF305426																		
24	594.4	45.0	61505	9	AF305426																		
25	594.4	45.0	61505	9	AF305426																		
26	594.4	45.0	61505	9	AF305426																		
27	594.4	45.0	61505	9	AF305426																		
28	594.4	45.0	61505	9	AF305426																		
29	594.4	45.0	61505	9	AF305426																		
30	594.4	45.0	61505	9	AF305426																		
31	594.4	45.0	61505	9	AF305426																		
32	594.4	45.0	61505	9	AF305426																		
33	594.4	45.0	61505	9	AF305426	</																	

C	11	493.6	37.4	693	6	AX392430	AX392430 Se
---	----	-------	------	-----	---	----------	-------------

13	417.8	31.6	812	6	AX011709	AX011709 Se
14	401.4	30.4	400	5	AX011707	AX011707 Se

C	15	362	27.4	156075	2	AL837509	AL837509	MU
---	----	-----	------	--------	---	----------	----------	----

17	350	26.5	2170	9	AK055028	AK055028 HC
----	-----	------	------	---	----------	-------------

19	349	26.4	8093	9	AF009426	AF009426	HC
----	-----	------	------	---	----------	----------	----

22	281	21.3	8039	9	AF009427	AF009427	HC
----	-----	------	------	---	----------	----------	----

24	250.2	18.9	173709	9	AP001010	AP001010 HC
----	-------	------	--------	---	----------	-------------

26	240.8	18.2	766	10	BC022716	BC022716
----	-------	------	-----	----	----------	----------

28	231	17.5	183619	2	AC111069	AC111069 Mu
----	-----	------	--------	---	----------	-------------

C	30	215.8	16.3	167758	2	AC110189	AC110189	Mu
---	----	-------	------	--------	---	----------	----------	----

32	120	9.1	159824	2	AC111878	AC111878	Re
32	120	9.1	159824	2	AC111878	AC111878	Re

[illegible]

36	99.6	7.5	108782	2	AC123508	AC123508	R
36	99.6	7.5	108782	2	AC123508	AC123508	R

38	99	7.5	83629	2	AC110485	AC110485	Re
----	----	-----	-------	---	----------	----------	----

40	96.4	7.3	154684	2	AC128630	AC128630 R
----	------	-----	--------	---	----------	------------

C	42	94.4	7.1	208757	2	AC094591	AC094591	Re
C	42	94.4	7.1	208757	2	AC094591	AC094591	Re

C 44	94	7.1	151283	2	AC107540	AC107540 Re
C 44	94	7.1	151283	2	AC107540	AC107540 Re

## ALIGNMENTS

AX392417

DEFINITION Sequence 1 from Patent WO0216416.

VERSION	AX392417.1	GI:19700732
---------	------------	-------------

SOURCE	human.
--------	--------

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele-

## REFERENCE 1

Turl, T.G.

**TITLE**      **Diagnosis and treatment of cardiovascular conditions**

JOURNAL Patent: WO 0216416-A 1 28-FEB-2002;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US); PFIZER INC. (US)

FEATURES  
Location/Qualifiers  
1. 1321

source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
413.1276  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD29005.1"  
/db\_xref="GI:19700733"

CDS  
/translation="MHRLMGVNSTAAAGAPNVCNCKRSLFOSMETELEFVQI  
IIVVMVMVAVITCLSHYKLSRSTFISHSGRRKREDALESGLMSESVGN  
GIPEPOYAPRPTDLAVPEPAQERFHRQPTPYLOHIDIDPITISLDEGEPP  
YOGCTLOLRPEOOLEINRESVRAPNRIIFDSDLMDSALHGGPCPPSSNSISATC  
YSGRBMGGPPPTSEVYGHYPGSSFOHOSSPPSLIEGRLHHTHIAPLESAIWS  
KEKDKOKHPL"

BASE COUNT 223 a 493 c 440 g 165 t  
ORIGIN  
Query Match 100.0%; Score 1321; DB 6; Length 1321;  
Best Local Similarity 100.0%; Pred. No. 9.5e-176;  
Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACGCGGTCTCGAGCAAAACCGATCTCTTGACTTGATAGAGAGAGAGGGG 60  
DB 1 CGACGCGGTCTCGAGCAAAACCGATCTCTTGACTTGATAGAGAGAGAGGGG 60  
QY 61 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
DB 61 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
QY 121 GCGGCGAGCG 180  
DB 121 GCGGCGAGCG 180  
QY 121 GCGGCGAGCG 180  
DB 121 GCGGCGAGCG 180  
QY 181 AGCG 240  
DB 181 AGCG 240  
QY 241 TGAGCG 300  
DB 241 TGAGCG 300  
QY 301 CTGCG 360  
DB 301 CTGCG 360  
QY 301 CTGCG 360  
DB 301 CTGCG 360  
QY 361 CG 420  
DB 361 CG 420  
QY 421 CTGATGAGGAGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
DB 421 CTGATGAGGAGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
QY 481 GTGCAAGTCAACAGCTCTTTGTTCCAGAGCATGAGATACAGAGCTGAGTTGTTCA 540  
DB 481 GTGCAAGTCAACAGCTCTTTGTTCCAGAGCATGAGATACAGAGCTGAGTTGTTCA 540  
QY 541 GATCATCATCATCTGT 600  
DB 541 GATCATCATCATCTGT 600  
QY 601 CCACTACAAGTGTCTGACAGGTCCTTATCAGCGGACAGCCAGGAGGAGAGAGA 660  
DB 601 CCACTACAAGTGTCTGACAGGTCCTTATCAGCGGACAGCCAGGAGGAGAGAGA 660  
QY 661 AGATGCCCTGTCTCAGAGAGATGCTGTGCGCTCGAGAGACAGAGTGTGAGCAACGG 720  
DB 661 AGATGCCCTGTCTCAGAGAGATGCTGTGCGCTCGAGAGACAGAGTGTGAGCAACGG 720  
QY 721 AATCCAGAGCGCGAGTGTCTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780

DB 721 AATCCAGAGCGCGAGTGTCTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780  
QY 781 CTGCG 840  
DB 781 CTGCG 840  
QY 841 GATGAGCTGCGACCGCCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
DB 841 GATGAGCTGCGACCGCCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
QY 901 CCCCTGACCCCTCCAGCTTGCGGAGCCCGAGCAGCACTGGAACCTAACCAGGAGTGGT 960  
DB 901 CCCCTGACCCCTCCAGCTTGCGGAGCCCGAGCAGCACTGGAACCTAACCAGGAGTGGT 960  
QY 961 GCGGCGACCG 1020  
DB 961 GCGGCGACCG 1020  
QY 1021 CGGCG 1080  
DB 1021 CGGCG 1080  
QY 1081 GCGCATGAGAGGCG 1140  
DB 1081 GCGCATGAGAGGCG 1140  
QY 1141 CTTCAGACACAGAGAGAGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200  
DB 1141 CTTCAGACACAGAGAGAGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200  
QY 1201 CACACACATCCG 1260  
DB 1201 CACACACATCCG 1260  
QY 1261 AGGACACCTCTCTTAGGTTCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320  
DB 1261 AGGACACCTCTCTTAGGTTCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320  
QY 1321 G 1321  
DB 1321 G 1321

RESULT 2  
AF305616 4839 bp mRNA linear PRI 21-OCT-2001  
LOCUS Homo sapiens STAG1/PMEP1 mRNA, complete cds.  
ACCESSION AF305616  
VERSION AF305616.1 GI:16303741  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 4839)  
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.  
Characterization of a novel gene, STAG1/PMEP1, upregulated in  
renal cell carcinoma and other solid tumors  
Mol. Carcinog. 32 (1), 44-53 (2001)  
JOURNAL  
MEDLINE  
21453682  
PUBMED  
11568975  
REFERENCE  
2 (bases 1 to 4839)  
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.  
Direct Submision  
Submitted (14-SEP-2000) Centre for Molecular Biotechnology, QLD  
Queensland University of Technology, 2 George St, Brisbane, QLD  
4001, Australia  
FEATURES  
source  
1.4839  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="20"

```

/mapp="20q13.2.q13.33"
321.1184
/notes="unknown function"
/codon_start=1
/product="STAG1/PMEPA1"
/protein_id="PALL16781.1"
/db_xref="GI:16303742"
/translation="MHRLMGVNSTAAALAAAGOPNVSCINCKRSLFQSMETLEFPVQI
IILVVMVAVPVITCLLSHKIKLSARFISHSQGRLEHLSSEGLIMBESIVSGN
GIPDPYVAPRPPDRPLAVPFPQREHFRHFRPDRHIDLPITLSIDGDEPP
YQSGCTQLDLPDEQGLINRESVAPRNRTFDEDLMDASALGCPSSNSGISATCC
YSGSGRMEGPPPTVSEVIGHYRQSSFQHQSQSPSLLEGRLHHTHAPLESANLMS
KERNKQGHPL"
polyA_signal      2158..2163
polyA_signal      2463..2468
polyA_signal      4818..4823
BASE COUNT       1103 a 1313 c 1290 g 1133 t
ORIGIN
Query Match      93.0%; Score 1229; DB 9; Length 4839;
Best Local Similarity 100.0%; Prid. No. 5.6e-163;
Matches 1229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 93 GGAAGCAGGCGGAGCGAGCGCTCAGCCCGGGGGCGGCGCGCCCGCTGCAGCCCAT 152
Db 1 GGAAGCTAGGGGCGAGAGGCTCAGCCCGGGGGCGGCGCGCCCGCTGCAGCCCAT 60
QY 153 TTCCGAGCGCCACCGCGGGGCACTCCCGACGCCCCCGGGGCTGCCGAGGGAGCGGG 212
Db 61 TTCCGAGCGCCACCGCGGGGCACTCCCGACGCCCCCGGGGCTGCCGAGGGAGCGGG 120
QY 213 GGGCGCAGCGAGCGGGTCCCGCGCACTGAGCCCGCGCGCGCCCGGGAACTTGGCGGC 272
Db 121 GGGCGCAGCGAGCGGGTCCCGCGCACTGAGCCCGCGCGCGCCCGGGAACTTGGCGGC 180
QY 273 GACCCGACCCCGGGGAGCGCGGGGGCGGCTCCCCCGCGCGGCTCTCGATGGGGGG 332
Db 181 GACCCGACCCCGGGGAGCGCGGGGGCGGCTCCCCCGCGCGGCTCTCGATGGGGGG 240
QY 333 CCCAGCTCGGGGAGCGCGGGCGCGGAGCCCCCGCGGCGGCCCGGAGCCCCCGCGCGC 392
Db 241 CCCAGCTCGGGGAGCGCGGGCGCGGAGCCCCCGCGGCGGCCCGGAGCCCCCGCGCGC 300
QY 393 GCCGCGCGCGCGCGCGCTCAGTCACCGCTTGATGGGGGTCAACAGCACCGCGCGCG 452
Db 301 GCCGCGCGCGCGCGCGCTCAGTCACCGCTTGATGGGGGTCAACAGCACCGCGCGCG 360
QY 453 CCGCGCGGCGACCCCAATGTCTCTCAGGTGCACTGCAAAAGCGCTTTGTTCCAGAGCA 512
Db 361 CCGCGCGGCGACCCCAATGTCTCTCAGGTGCACTGCAAAAGCGCTTTGTTCCAGAGCA 420
QY 513 TGGAGATCAGGAGCGAGTGGAGTTTGTTCAGATCATCATCGTGGTGGATGATGATGA 572
Db 421 TGGAGATCAGGAGCGAGTGGAGTTTGTTCAGATCATCATCGTGGTGGATGATGATGA 480
QY 573 TGGTGGTGGTGAATCACGTGCTGTGAGCACATACAACTGTCTCAGCGTCCATCA 632
Db 481 TGGTGGTGGTGAATCACGTGCTGTGAGCACATACAACTGTCTCAGCGTCCATCA 540
QY 633 GCCGCGCAGCGCAGGGGGGGGAGAGAAAGATGCCCTGTCTCAGAAAGATGCTGTGGC 692
Db 541 GCCGCGCAGCGCAGGGGGGGGAGAGAAAGATGCCCTGTCTCAGAAAGATGCTGTGGC 600
QY 693 CCTCGAGAGGACAGATGTCTCAGGCAACGGAATCCAGAGGCGGAGGCTACAGCGCGCTC 752
Db 601 CCTCGAGAGGACAGATGTCTCAGGCAACGGAATCCAGAGGCGGAGGCTACAGCGCGCTC 660
QY 753 GGGCCACCGACCGCTGGCGCGCTGCGCGCTTCGCCACGCGGGAGCGCTTCCACGCGCTTC 812
Db 661 GGGCCACCGACCGCTGGCGCGCTGCGCGCTTCGCCACGCGGGAGCGCGCTTCCACGCGCTTC 720
QY 813 AGCCCACTATTCGACTGCGAGCACGAGATCGACCTGCACCCACATCTCGCTGTCA 872

```

Db	721	AGCACCCTATTCGACCTGCAGCACGAGATCGACTTCCACCAACCATCTTCGTGTACG	780
Oy	873	ACGGGAGAGAGCCCCCACCCTTACACAGGGCCCCCTGCACCCCTCAGCTTGSGGAGCCCGAGC	932
Db	781	ACGGGAGAGAGAGCCCCCACCCTTACACAGGGCCCCCTGCACCCCTCAGCTTGSGGAGCCCGAGC	840
Oy	933	AGCAGCTGGAACTGAACCCGGGAGTGTGGTGGCGGACCCCCCAAACGAAACATCTTTCAGCA	992
Db	841	AGCAGCTGGAACTGAACCCGGGAGTGTGGTGGCGGACCCCCCAAACGAAACATCTTTCAGCA	900
Oy	993	GTCAGCTGATGATGTGTCCACAGGCTGGGCGGCCCTTGCCTCCCGCCAGTAGTAATCGGGACA	1052
Db	901	GTGACCTATGATGTGTGTGTCCACAGGCTGGGCGGCCCTTGCCTCCCGCCAGTAGTAATCGGGACA	960
Oy	1053	TCAAGCCACAGCTGTCTACAGGCAAGCCGGCGGCATGTGAAGGGGCGCCGCCCACTTACAGCG	1112
Db	961	TCAGGCGCCACGTGTCTACAGGCAAGCCGGCGGCATGTGAAGGGGCGCCGCCCACTTACAGCG	1020
Oy	1113	AGGTATATGGGCACTACCCGGGGGTCTCTCTTCAGACACCAGCAGACAGTGGGCCGCCCT	1172
Db	1021	AGGTATATGGGCACTACCCGGGGGTCTCTCTTCAGACACCAGCAGACAGTGGGCCGCCCT	1080
Oy	1173	CCTTCTGAGAGGGGAGCCGGGGTCCACACACACATGCGGCCCTTACAGAGCGCAGCCA	1232
Db	1081	CCTTCTGAGAGGGGAGCCGGGGTCCACACACACATGCGGCCCTTACAGAGCGCAGCCA	1140
Oy	1233	TCTGGAGCAAGAAGAGATAAACAAGAACACCCCTCTCTAGAGGTCCCGCAGGGGGGCC	1292
Db	1141	TCTGGAGCAAGAAGAGATAAACAAGAACACCCCTCTCTAGAGGTCCCGCAGGGGGGCC	1200
Oy	1293	GGGCTGGGGCTGCTAGGTGAAGAAGCGCAG	1321
Db	1201	GGGCTGGGGCTGCTAGGTGAAGAAGCGCAG	1229
RESULT 3			
BC015918			
LOCUS	BC015918	1061 bp	mRNA linear PRI 17-OCT-2001
DEFINITION	Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.		
ACCESSION	BC015918		
VERSION	BC015918.1 GI:16198474		
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eumaxyotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Strausberg,R.		
JOURNAL	Direct Submissions Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk Email: cgaabbs@email.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guo, Letticia Hsiao, Martin Krywinski, Reta Kusche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Matthews, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stoltz, Michael Thorne, Miraneda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov





```

QY 653 AGGAGAGAAAGATGCGCTGCTCCTCAGAGAGATGCTGCGCTCGAGAGACAGATGTCA 712
Db 241 AGGAGAGAAAGATGCGCTGCTCCTCAGAGAGATGCTGCGCTCGAGAGACAGATGTCA 300
QY 713 GGCAGACGAAATCCAGAGCGCAGGTCTACGCCCGCTCGGCCACCGACCGCTGGCC 772
Db 301 GGCAGACGAAATCCAGAGCGCAGGTCTACGCCCGCTCGGCCACCGACCGCTGGCC 360
QY 773 GTGCGCGCTTCCGCCAGCGGAGCGCTTCACCGCTTCACGCCACCTTCCGCTACCTG 832
Db 361 GTGCGCGCTTCCGCCAGCGGAGCGCTTCACCGCTTCACGCCACCTTCCGCTACCTG 420
QY 833 CAGCAGAGATGACCTGCGCCACCGACCATCTGCTGTCAGAGCGGAGGAGCGCCGACCC 892
Db 421 CAGCAGAGATGACCTGCGCCACCGACCATCTGCTGTCAGAGCGGAGGAGCGCCGACCC 480
QY 893 TACCGAGCGCGCTGCGACCTTCAGCTTCGCGGAGCGCCGAGCAGCAGTGAATGAACCG 952
Db 481 TACCGAGCGCGCTGCGACCTTCAGCTTCGCGGAGCGCCGAGCAGCAGTGAATGAACCG 540
QY 953 GAGTCGGTGGCGGACCGCCCAACAGACCATCTTTCAGCAGTGAATGATGATGCTCC 1012
Db 541 GAGTCGGTGGCGGACCGCCCAACAGACCATCTTTCAGCAGTGAATGATGATGCTCC 600
QY 1013 AGGCTGGGCGCGCTGCGCCCGCCGAGCAGTGAATGATGATGATGATGATGATGATG 1072
Db 601 AGGCTGGGCGCGCTGCGCCCGCCGAGCAGTGAATGATGATGATGATGATGATGATG 660
QY 1073 AGCGGCGGCGCGATGAGAGGCGCGCGCCGACCTTACAGCAGTGAATGATGATGATG 1132
Db 661 AGCGGCGGCGCGATGAGAGGCGCGCGCCGACCTTACAGCAGTGAATGATGATGATG 720
QY 1133 GGGTCTCTCTCCAGCAGCAGCAGAGAGTGGGCGCGCTCTCTCTCTCTCTCTCTCT 1192
Db 721 GGGTCTCTCTCCAGCAGCAGCAGAGAGTGGGCGCGCTCTCTCTCTCTCTCTCTCT 780
QY 1193 CTCACACACACACATCGCGCCCTGAGAGAGCGCAGCAGTCTGAGCAGCAAGAGAT 1252
Db 781 CTCACACACACACATCGCGCCCTGAGAGAGCGCAGCAGTCTGAGCAGCAAGAGAT 840
QY 1253 AAACAGAAAGACACCTCTC 1273
Db 841 AAACAGAAAGACACCTCTC 861

RESULT 5
AF224278 1141 bp mRNA linear PRI 18-JUL-2000
LOCUS Homo sapiens PMEPA1 protein (PMEPA1) mRNA, complete cds.
DEFINITION AF224278
ACCESSION AF224278
VERSION AF224278.1 GI:9255808
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1141)
XU,L., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
Moul,J.W., and Srivastava,S.
A novel androgen-regulated gene, PMEPA1, located on chromosome
20q13 exhibits high level expression in prostate
Genomics 66 (3), 257-263 (2000)
TITLE
JOURNAL MEDLINE
PUBMED 20334621
10873380
REFERENCE
AUTHORS 2 (bases 1 to 1141)
XU,L., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
Moul,J.W., and Srivastava,S.
Direct Submission
Submitted (12-JAN-2000) CPDR, USUHS, 1530 East Jefferson Street,
Rockville, MD 20852, USA
FEATURES
source
1. 1141
/organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.31-13.33"
/cell_line="LNCap"
1. 1141
/gene="PMEPA1"
96..854
CDS
/gene="PMEPA1"
/note="type 1b transmembrane protein; expression is
induced in response to the synthetic androgen, R1881;
expression is abundant in, and restricted to, prostate
glandular epithelial cells; similar to the predicted
protein encoded by sequence deposited at Genbank Accession
Number NP_004329"
/codon_start=1
/product="PMEPA1 protein"
/db_xref="GI:9255809"
/translation="MALEPFOVQIIIVVMMVVVITCLLSHYKLSARSFISRHSG
RRREDLSSBGLMPSESVYSGNGTPEPOVYAPRRRTDLAVPPRFRFRFPQTY
PLYQHEIDLPPTISLSDGSEPPYQPTQLQDLPDQLELNRESYRAPRNRITUSD
LMDSARLGPSPSSNSGISAIVCYGSGRMEGPPYSEVIGHYPPSSPFOHDOSSGP
SLEGTRLHHTHTAPLESAAIWSKERDKKHPL"
BASE COUNT 271 a 350 c 336 g 184 t
ORIGIN
Query Match 60.7%; Score 802.2; DB 9; Length 1141;
Best Local Similarity 98.4%; Pred. No. 3.7e-103;
Matches 810; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 499 TTGTGTCAGAGATGAGATACAGGAGCGAGTGGTGTTCAGATCATCATCGTGTGT 558
Db 77 TCTCTCGGAACAGGAGCATAGCGAGCTGAGCTTGTTCAGATCATCATCGTGTGT 136
QY 559 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 618
Db 137 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 196
QY 619 ACGGTCTTATATACGCGGACAGCCAGGGGCGGAGAGAGAGATGCTCTCTCTCTCT 678
Db 197 ACGGTCTTATATACGCGGACAGCCAGGGGCGGAGAGAGAGATGCTCTCTCTCTCT 256
QY 679 AGGATGCTGTGGCGCTTCGGAGACACAGTGTGAGCAGCAAGGATCCAGAGCGCAGGT 738
Db 257 AGGATGCTGTGGCGCTTCGGAGACACAGTGTGAGCAGCAAGGATCCAGAGCGCAGGT 316
QY 739 CTAGCGCCCGCTTCGGCCACCGACCGCTGCGCGCTGCGCCCTTCGCGCCAGCGGAGCG 798
Db 317 CTAGCGCCCGCTTCGGCCACCGACCGCTGCGCGCTGCGCCCTTCGCGCCAGCGGAGCG 376
QY 799 CTTCACCGCTTCGAGCGCCACCTATCTGCTGCTGACACAGATGCACTGCCACCCAC 838
Db 377 CTTCACCGCTTCGAGCGCCACCTATCTGCTGCTGACACAGATGCACTGCCACCCAC 436
QY 859 CATCTGCTGTGAGAGCGGAGAGAGCGCCACCTTACAGAGCGCCCTGCAACCTTCAGCT 918
Db 437 CATCTGCTGTGAGAGCGGAGAGAGCGCCACCTTACAGAGCGCCCTGCAACCTTCAGCT 496
QY 919 TCGGAGCCCGAGAGCAGCTGGAACCTGAGAGCGGAGTGGTGGCGCACCCCAACAG 978
Db 497 TCGGAGCCCGAGAGCAGCTGGAACCTGAGAGCGGAGTGGTGGCGCACCCCAACAG 556
QY 979 AACCATTTTGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATG 1038
Db 557 AACCATTTTGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATG 616
QY 1039 CAGTAACTCGGAGCATGCGCACGCTGCTAGCGAGCGGCGGCGCATGAGAGGCGCGCC 1098
Db 617 CAGTAACTCGGAGCATGCGCACGCTGCTAGCGAGCGGCGGCGCATGAGAGGCGCGCC 676
QY 1099 GCCCAGCTACAGGAGGTATCGGCCACTACCGCGGGTCTCTCTTCCAGCAGCAGAG 1158
Db 677 GCCCAGCTACAGGAGGTATCGGCCACTACCGCGGGTCTCTCTTCCAGCAGCAGAG 736

```

	Query Match	60.6%	Score 800	DB 9	Length 1818
	Best Local Similarity	100.0%	Pred. No. 6.9e-103		
Matches	800	Conservative	0	Mismatches	0; Indels 0; Gaps 0;
Y	522 CGAGCTGAGATTGTTCATCATCATCATCGTGGTGATGATGGTGGTTGG				581
b	104 CCGAGCGCGAGCTTTGTTCCAGATCATCATCATCGTGGTGATGATGGTGG				163
Y	582 TGATCAGCTCCCTGCAGACACTCAAGCTGTGACAGGCTCTTCATCAGCCGCACA				641
b	164 TGATCAGCTCCCTGCAGACACTCAAGCTGTGACAGGCTCTTCATCAGCCGCACA				223
Y	642 GCCAGGGCGCGAGGAGAAAGATGCCCTGTCTCGAAGAATGCTGTGGCCCTCGGAGA				701
b	224 GCCAGGGCGCGAGGAGAAAGATGCCCTGTCTCGAAGAATGCTGTGGCCCTCGGAGA				283
Y	702 GCACAGTGTCAAGCAACGAATCCAGACGCGCAGGTTAGCCCCGGCCTGGGCCACCG				761
b	284 GCACAGTGTCAAGCAACGAATCCAGACGCGCAGGTITAGCCCCGGCCTGGGCCACCG				343
Y	762 ACCGCGTGGCCCTGGCCGCTTCGCCCAAGCGGAGCGCTTCACACCGCTTCACAGCCACCT				821

```

/organism="Mus musculus"
/db_xref="taxon:10090"
20. .844
/feature="unamed protein product"
/codon_start=1
/protein_id="CAD29011.1"
/db_xref="gi:19700745"
/translation="MGVNGTAAAAAGQPVNVCACNCQSLPFSMEITELFEYQIVVIV
VMVAQVVMVITLASHKLSARFISRHQAARRDGLSSECLMWSESTVSGMPE
QVAAPPEPTRLAAPPFIQNRSPQYPIQLQHEIALPTISLDSGEAPPYQGPCTLO
LRPEQDQLNRESVAPNRKTIFFSDLLIDSTMLGGPCPPSSNGISATCVSSGRME
GPPTTSEVYGHVPGSSFQHQSGSSLLSEGRRLHSHRIAPLENKEKQKQHPPL"
179 a 305 c 247 g 147 t

```



QY 1138 CTCCTTCAGACACAGAGAGAGTGGGGCCCTCTTCTGCTGAGGGAGCCCGCTCCA 1197  
 |||||  
 Db 57712 CTCCTTCAGACACAGAGAGAGTGGGGCCCTCTTCTGCTGAGGGAGCCCGCTCCA 57771  
 QY 1198 CCACACACATCGCGCCCTAGAGAGCGACCATCTGAGCAAGAGAGATTAACA 1257  
 |||||  
 Db 57772 CCACACACATCGCGCCCTAGAGAGCGACCATCTGAGCAAGAGAGATTAACA 57831  
 |||||  
 QY 1258 GAAAGAGACCTCTCTAGAGGTCCCGAGGGGGCCGGGCTGGGCTGAGTGAAGAAG 1317  
 |||||  
 Db 57832 GAAAGAGACCTCTCTAGAGGTCCCGAGGGGGCCGGGCTGGGCTGAGTGAAGAAG 57891  
 |||||  
 QY 1318 GCAG 1321  
 |||||  
 Db 57892 GCAG 57895

RESULT 9  
 HS718J7/c  
 DEFINITION  
 HS718J7 130435 bp DNA linear PRI 24-FEB-2001  
 Human DNA sequence from clone RP4-718J7 on chromosome 20q13.31-13.33 contains the PKC1 gene for soluble phosphoenolpyruvate carboxykinase 1, part of a novel gene similar to mouse DLM-1 (tumour stroma and activated macrophage protein), the 3' end of the TMEPAI gene encoding an androgen induced 1b transmembrane protein (PMEPAI), two putative novel genes, a CPG island, ESTs, STS and GSSs, complete sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 AL035541.15 GI:11546043  
 HNG: CPG Island; DLM-1; macrophage protein; PKC1; phosphoenolpyruvate carboxykinase; PMEPAI; transmembrane protein.

SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Sehra, H.  
 Direct Submission  
 Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 Requests: clonerequest@sanger.ac.uk  
 On Dec 5, 2000 this sequence version replaced gi:10198628.

## COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission only a small overlap as described above, as we submit sequences with the following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>  
 This sequence is the entire insert of clone RP4-718J7 the true left right end of clone RP5-1007E6 is at 71437 in this sequence. The true right end of clone RP4-579F20 is at 43945 in this sequence.  
 RP4-718J7 is from the library RPCI-4 constructed by the group of Plietier de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCYPAC2  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## FEATURES

## source

Location/Qualifiers  
 1..130435  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="q13.31-13.33"  
 /clone="RP4-718J7"  
 /clone\_lib="RPCI-4"  
 319..478  
 /note="5 copies 32 mer 67% conserved"  
 370..463  
 /note="2 copies 47 mer 87% conserved"  
 427..626  
 /note="4 copies 50 mer 71% conserved"  
 707..812  
 /note="53 copies 2 mer cc 61% conserved"  
 967..2752  
 /note="893 copies 2 mer gg 54% conserved"  
 982..2730  
 /note="33 copies 53 mer 54% conserved"  
 1177..1820  
 /note="4 copies 161 mer 64% conserved"  
 1204..1371  
 /note="3 copies 56 mer 75% conserved"  
 1358..2757  
 /note="28 copies 50 mer 54% conserved"  
 1675..2718  
 /note="18 copies 58 mer 55% conserved"  
 1843..2345  
 /note="9 copies 56 mer 64% conserved"  
 1852..2079  
 /note="4 copies 57 mer 86% conserved"  
 2028..2632  
 /note="11 copies 55 mer 60% conserved"  
 2087..2257  
 /note="3 copies 57 mer 83% conserved"  
 2228..2497  
 /note="5 copies 54 mer 75% conserved"  
 2583..2716  
 /note="2 copies 67 mer 82% conserved"  
 3378..3426  
 /note="LIMB4 repeat: matches 6088..6136 of consensus"  
 3799..3896  
 /note="Charliel repeat: matches 681..781 of consensus"  
 5331..5793  
 /note="WTR1B repeat: matches 14..466 of consensus"  
 5797..5988  
 /note="6 copies 32 mer 86% conserved"  
 5855..5978  
 /note="31 copies 4 mer gcac 61% conserved"  
 6474..6591  
 /note="MFR repeat: matches 91..218 of consensus"  
 6592..6723  
 /note="FLAM\_C repeat: matches 1..132 of consensus"  
 6952..7021  
 /note="L2 repeat: matches 2637..2705 of consensus"  
 7358..7671  
 /note="AlusX repeat: matches 1..312 of consensus"  
 8521..8554  
 /note="Alu repeat: matches 1..34 of consensus"  
 8863..9217  
 /note="MUTID repeat: matches 105..505 of consensus"  
 9228..9746  
 /note="L2 repeat: matches 1459..2028 of consensus"  
 9880..10089  
 /note="L2 repeat: matches 2292..2511 of consensus"  
 10094..10206  
 /note="Aluub repeat: matches 188..300 of consensus"  
 10263..10573  
 /note="Aluub repeat: matches 1..308 of consensus"  
 10658..10753  
 /note="L2 repeat: matches 2647..2749 of consensus"  
 11816..12380

```
/note="LTR19B repeat: matches 1. .580 of consensus"
12519. .12813
repeat_region
/note="L1MD2 repeat: matches 6032. .6331 of consensus"
13559. .13726
repeat_region
/note="3 copies 56 mer 73% conserved"
13585. .13716
repeat_region
/note="22 copies 6 mer ctctct 67% conserved"
13586. .13717
repeat_region
/note="66 copies 2 mer tc 68% conserved"
13588. .13715
repeat_region
/note="4 copies 32 mer 78% conserved"
13589. .13724
repeat_region
/note="34 copies 4 mer cttt 77% conserved"
complement(13622. .14142)
/note="match: GSS: Em:AQ592603"
14313. .14437
repeat_region
/note="LTR16C repeat: matches 257. .386 of consensus"
15015. .15110
repeat_region
/note="3 copies 32 mer 79% conserved"
15021. .15110
repeat_region
/note="45 copies 2 mer ac 80% conserved"
15026. .15109
repeat_region
/note="7 copies 12 mer 82% conserved"
15028. .15111
repeat_region
/note="14 copies 6 mer cacaca 82% conserved"
15029. .15108
repeat_region
/note="20 copies 4 mer acac 83% conserved"
15273. .15399
repeat_region
/note="L2 repeat: matches 2159. .2285 of consensus"
15420. .15721
repeat_region
/note="AluX repeat: matches 1. .303 of consensus"
16333. .16396
repeat_region
/note="2 copies 32 mer 98% conserved"
16575. .16690
repeat_region
/note="L2 repeat: matches 2596. .2711 of consensus"
18049. .18169
repeat_region
/note="MIR repeat: matches 86. .211 of consensus"
18312. .18438
repeat_region
/note="MIR repeat: matches 138. .250 of consensus"
18585. .18776
repeat_region
/note="MER20 repeat: matches 7. .217 of consensus"
19211. .19290
repeat_region
/note="MIR repeat: matches 72. .154 of consensus"
20413. .20463
repeat_region
/note="MIR repeat: matches 90. .143 of consensus"
complement(21005. .21672)
/note="match: GSS: Em:AQ748384"
complement(21187. .21672)
/note="match: GSS: Em:AQ776209"
complement(21278. .21663)
/note="match: GSS: Em:AQ136459"
complement(21392. .21672)
/note="match: GSS: Em:AQ892114"
21667. .21775
repeat_region
/note="MIR repeat: matches 73. .192 of consensus"
23577. .23646
repeat_region
/note="2 copies 35 mer 100% conserved"
24257. .24591
repeat_region
/note="L2 repeat: matches 2336. .2710 of consensus"
25580. .25611
repeat_region
/note="16 copies 2 mer tc 87% conserved"
26333. .26643
repeat_region
/note="Alu repeat: matches 1. .310 of consensus"
27603. .27684
repeat_region
/note="L2 repeat: matches 2661. .2739 of consensus"
27823. .28041
repeat_region
/note="MIR repeat: matches 38. .242 of consensus"
28129. .28259
repeat_region
/note="F1AM_C repeat: matches 1. .127 of consensus"
31203. .31356
repeat_region
/note="L1MB5 repeat: matches 6015. .6176 of consensus"
32224. .32522
repeat_region
/note="AluX repeat: matches 1. .298 of consensus"
```

```
repeat_region 33285. .33481
Query Match 45.0%; Score 594.4; DB 9; Length 130435;
Best Local Similarity 99.0%; Pred. No. 1.9e-74;
Matches 598; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 718 CGGAATCCAGAGCCGAGGCTCAGCCGCCGCGCCAGCAGACCGCCGCGGCTGCC 777
| 111 |
Db 128454 CTGCTTCCTCCAGCCGAGGTCTACGCCGCCGCTCGGCCACGACCGCCCTGGCGTGCC 128395
QY 778 GCCCTTGCCAGCGGAGCGCTTCCACCGCTTCCAGCCCGACCTATCCGATCGACGA 837
| 111 |
Db 128394 GCCCTTGCCAGCGGAGCGGCTTCCACCGCTTCCAGCCCGACCTATCCGATCGACGA 128335
QY 838 CGAGATGACCTGCGACCCACCATCTGCTGTACAGACGGGAGAGGCCGCCACCTTACCA 897
| 111 |
Db 128334 CGAGATGACCTGCGACCCACCATCTGCTGTACAGACGGGAGAGGCCGCCACCTTACCA 128275
QY 898 GGGCCCTGACCCCTCAGCTGCGGAGCCCGGACGACGACGCTGGAACCGGGAGTC 957
| 111 |
Db 128274 GGGCCCTGACCCCTCAGCTGCGGAGCCCGGACGACGACGCTGGAACCGGGAGTC 128215
QY 958 GGTGCGGCAACCCCAACAGAACCATCTTGACAGTGAACCTGATGATGTCAGAGCT 1017
| 111 |
Db 128214 GGTGCGGCAACCCCAACAGAACCATCTTGACAGTGAACCTGATGATGTCAGAGCT 128155
QY 1018 GGGCGGCCCTGCCGCCCGCCAGCACTAGTGGGATAGGCGCCAGTCTACGACCGG 1077
| 111 |
Db 128154 GGGCGGCCCTGCCGCCCGCCAGCACTAGTGGGATAGGCGCCAGTCTACGACCGG 128095
QY 1078 CGGCGCATGAGAGGGGGCGGCCCGCCACCTAGAGAGTATGAGGACACCGGGGTC 1137
| 111 |
Db 128094 CGGCGCATGAGAGGGGGCGGCCCGCCACCTAGAGAGTATGAGGACACCGGGGTC 128035
QY 1138 CTCTTCACGACACGAGAGAGAGTGGCGGCCCTCTCTGTGAGGGGAGCCCGCTCA 1197
| 111 |
Db 128034 CTCTTCACGACACGAGAGAGAGTGGCGGCCCTCTCTGTGAGGGGAGCCCGCTCA 127975
QY 1198 CCACACACATGCGCGCCCTAGAGAGCGACCATCTGAGCAAGAAGAGATTAAACA 1257
| 111 |
Db 127974 CCACACACATGCGCGCCCTAGAGAGCGACCATCTGAGCAAGAAGAGATTAAACA 127915
QY 1258 GAAAGACACCCCTCTAGGGTCCCAAGGGGGCGGGGCTGCGGCTCGTAGGGAAGA 1317
| 111 |
Db 127914 GAAAGACACCCCTCTAGGGTCCCAAGGGGGCGGGGCTGCGGCTCGTAGGGAAGA 127855
QY 1318 GCAG 1321
| 111 |
Db 127854 GCAG 127851
RESULT 10
HSJ1059L7/c 150224 bp DNA linear PRI 24-FEB-2001
LOCUS
DEFINITION
Human DNA sequence from clone RS5-105917 on chromosome
20q13.2-13.33 Contains the 5' end of the TMERAI (PMEPAI) gene
encoding an androgen induced 1b transmembrane protein, ESTs, STSS,
GSSs and two CpG islands, complete sequence.
ACCESSION
AL121913.4 GI:7161781
VERSION
AL121913.4
KEYWORDS
HTG; CpG island; PMEPAI; TMERAI; transmembrane protein.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 150224)
AUTHORS
Skuce,C.
TITLE
Direct Submission
JOURNAL
Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Mar 6, 2000 this sequence version replaced gi:7007305.
COMMENT
During sequence assembly data is compared from overlapping clones.
```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences without a small overlap as described above.

Only a small overlap between sequences is used to associate primary accession numbers given in the feature table with their source databases: Em1, EMBL, SW, SWISSPROT, Tr1, TREMBL, Wp1, WORMEP; Information on the WORMEP database can be found at [http://www.sanger.ac.uk/projects/C\\_elegans/wormep](http://www.sanger.ac.uk/projects/C_elegans/wormep) This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HG/Chr20>

This sequence is the entire insert of clone RP5-105917. The true left end of clone RP11-402P1 is at 106677 in this sequence. All regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats. All regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-105917 is from the library RCR1-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pcvpac2.

```

FEATURES
Source
Location/Qualifiers
1..150224
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="q13.2-13.33"
/clone="RP5-1059L7"
/clone_1lb="RPC1-5"
/clone_355
repeat_region
283..355
/note="MIR repeat: matches 71..145 of consensus"
1509..1628
/note="MIR repeat: matches 46..165 of consensus"
repeat_region
2076..2155
/note="L2 repeat: matches 2616..2707 of consensus"
gene
complement(2445..53425)
/gene="TMEPA1"
complement(join(<2445..2599,53362..53425))
/gene="TMEPA1"
/mrna
/product="dJ1059L7.1.2 (androgen induced type 1b transmembrane protein (PMEPA1), isoform 2)"
/note="match: cDNAs: Em:AF242478
match: ESTs: Em:AA088767"
evidence="not_experimental"
complement(join(<2445..2599,53362..53365))
/gene="TMEPA1"
/cds
/note="continues in dJ718J7 (AL035541)"
/codon_start=1
evidence="not_experimental"
/product="dJ1059L7.1.2 (androgen induced type 1b transmembrane protein (PMEPA1), isoform 2)"
/protein_id="CAC32857.1"
/db_xref="GI:13160408"
/translation="MALEFVQIILIVVMVYITCLSHYKLSRSPISRHSGC
RRREDLSS"
complement(join(<2445..2599,52376..52477))
/gene="TMEPA1"
/product="dJ1059L7.1.1 (androgen induced type 1b transmembrane protein (PMEPA1), isoform 1)"
/note="match: cDNAs: Em:AF009426 Em:AF009427 Em:AF009425
Em:AF009424
Em:AF009429 Em:AA249792 Em:AI594390 Em:AA128075
match: ESTs: Em:AA088767"
evidence="not_experimental"
complement(join(<2445..2599,52376..52472))
/gene="TMEPA1"
/note="continues in dJ718J7 (AL035541)"

```

```

match: proteins: Tr:O15166 Tr:O15168"
/codon_start=1
/evidence=not_experimental
/product="drl059L7.1.1 (androgen induced type 1b
transmembrane protein (PMEPA1), isoform 1)"
/protein_id="CA88144.1"
/db_xref="GI:7619746"
/db_xref="STRDBL:Q2NT9"
/translation="MGVNSTAAAGAPNVCNCNKRSLFQSMETLEFVQIIITV
VNMVMVVITCLLSHYKLSRSLFSLRISGRRREDALSS"
repeat_region
  /note="2538..2567
  /note="10 copies 3 mer cac 90% conserved"
misc_feature
  /note="3128..3606
  /note="match: GSS: Em:A059507"
repeat_region
  /note="3214..3324
  /note="3 copies 37 mer 83% conserved"
repeat_region
  /note="3339..3413
  /note="3 copies 25 mer 92% conserved"
repeat_region
  /note="3390..3921
  /note="19 copies 28 mer 61% conserved"
repeat_region
  /note="3394..3893
  /note="25 copies 20 mer 59% conserved"
repeat_region
  /note="3407..3892
  /note="27 copies 18 mer 58% conserved"
repeat_region
  /note="3421..3910
  /note="7 copies 70 mer 62% conserved"
repeat_region
  /note="3473..3904
  /note="12 copies 36 mer 58% conserved"
repeat_region
  /note="3533..3932
  /note="10 copies 38 mer 64% conserved"
repeat_region
  /note="3573..3697
  /note="5 copies 25 mer 76% conserved"
repeat_region
  /note="3746..3921
  /note="8 copies 22 mer 64% conserved"
repeat_region
  /note="4032..4069
  /note="19 copies 2 mer tt 81% conserved"
repeat_region
  /note="4792..5001
  /note="3 copies 70 mer 82% conserved"
repeat_region
  /note="5002..5190
  /note="MER20 repeat: matches 1..200 of consensus"
repeat_region
  /note="5224..5282
  /note="MIR repeat: matches 94..154 of consensus"
repeat_region
  /note="8103..8205
  /note="MIR repeat: matches 48..161 of consensus"
misc_feature
  /note="complement(8658..9277)
  /gene="TMEPA1"
  /note="match: GSS: Em:A0881699"
repeat_region
  /note="9224..9382
  /note="MIR repeat: matches 81..239 of consensus"
misc_feature
  /note="9286..9706
  /note="match: GSS: Em:A0565377"
misc_feature
  /note="9295..9861
  /note="match: GSS: Em:A0757992"
repeat_region
  /note="10035..10140
  /note="Charli4 repeat: matches 39..146 of consensus"
repeat_region
  /note="11175..11343
  /note="MIR repeat: matches 11..196 of consensus"
repeat_region
  /note="11630..12037
  /note="L2 repeat: matches 2290..2705 of consensus"
repeat_region
  /note="12054..12362
  /note="AlusX repeat: matches 1..311 of consensus"
repeat_region
  /note="14448..14493
  /note="3 copies 2 mer at 76% conserved"
repeat_region
  /note="14634..14813
  /note="MER20 repeat: matches 1..187 of consensus"
misc_feature
  /note="15096..15497
  /note="Cpg island"
  /note="evidence=not_experimental"
repeat_region
  /note="15690..15935
  /note="MIR repeat: matches 7..254 of consensus"
repeat_region
  /note="16989..17387
  /note="MIR214 repeat: matches 1..374 of consensus"
repeat_region
  /note="17647..17703

```

misc\_feature  
/note="L2 repeat: matches 2679. .2736 of consensus"  
17699. .18179  
/note="match: GSS: Em:AQ703107"  
17902. 18039  
repeat\_region  
/note="MIR repeat: matches 107. .243 of consensus"  
19070. .19291  
/note="MIR repeat: matches 35. .261 of consensus"  
19954. .20005  
repeat\_region  
/note="L2 repeat: matches 2358. .2411 of consensus"  
20264. .20432  
/note="L2 repeat: matches 2581. .2750 of consensus"  
20530. .21282  
/note="L2 repeat: matches 5595. .6341 of consensus"  
21980. .22192  
/note="L2 repeat: matches 2322. .2562 of consensus"  
23102. .23203  
repeat\_region  
/note="MIR repeat: matches 117. .218 of consensus"  
23231. .23438  
/note="MER20 repeat: matches 3. .218 of consensus"  
23439. .23512  
repeat\_region  
/note="MIR repeat: matches 182. .257 of consensus"  
24460. .24580  
/note="L2 repeat: matches 6102. .6221 of consensus"  
24752. .25042  
/note="L2 repeat: matches 2386. .2709 of consensus"  
25208. .25295  
repeat\_region  
/note="MIR repeat: matches 61. .155 of consensus"  
26450. .26485  
repeat\_region  
/note="18 copies 2 mer by 97% conserved"  
29374. .29501  
repeat\_region  
/note="MIR repeat: matches 55. .190 of consensus"  
29550. .29728  
repeat\_region  
/note="MIR repeat: matches 69. .248 of consensus"  
30974. .31029

Query Match 39.6%; Score 522.8; DB 9; Length 150224;  
Best Local Similarity 99.6%; Pred. No. 1.9e-64;  
Matches 524; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGACCCGGGCTCTCGAGAGCAACCCGATCTCTTGACTTAATAGAGAGAGAGCGG 60  
Db 52896 CGACCCGGGCTCTCGAGAGCAACCCGATCTCTTGACTTAATAGAGAGAGAGCGG 52837  
QY 61 CGCGGG 120  
Db 52836 CGCGGG 52777  
QY 121 GCGCGAGCG 180  
Db 52776 GCGCGAGCG 52717  
QY 181 ACGCCCGGGGGGCTGCGAGAGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 240  
Db 52716 ACGCCCGGGGGGCTGCGAGAGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGG 52657  
QY 241 TGAGCG 300  
Db 52656 TGAGCG 52597  
QY 301 CTCCCG 360  
Db 52596 CTCCCG 52537  
QY 361 CCCCCGG 420  
Db 52536 CCCCCGG 52477  
QY 421 CTTGATGGGGGTCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
Db 52476 CTTGATGGGGGTCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 52417  
QY 481 GTGCAACTGCAAAAGCTCTTTGTTCCAGAGCATGAGATCAAGGAG 526  
|||||

Db 52416 GTGCAACTGCAAAAGCTCTTTGTTCCAGAGCATGAGATCAAGGAG 52371  
RESULT 11  
AX392430/c 693 bp DNA linear PAT 23-MAR-2002  
LOCUS  
DEFINITION  
Sequence 14 from Patent WO0216416.  
ACCESSION  
AX392430  
VERSION  
AX392430.1 GI:19700746  
KEYWORDS  
human.  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and  
Turli,T.G.  
TITLE  
Diagnosis and treatment of cardiovascular conditions  
JOURNAL  
Patent: WO 0216416-A 14 28-FEB-2002;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)  
FEATURES  
source  
1. 693  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 101 a 205 c 237 g 149 t 1 others  
ORIGIN

Query Match 37.4%; Score 493.6; DB 6; Length 693;  
Best Local Similarity 97.4%; Pred. No. 6e-60;  
Matches 533; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 776 CCGGCTTCGGCCAGCGGGAGCGCTTCACCGCTTCACCGCCACCTAATCCGTACCTGAC 835  
Db 693 CCGGCTTCGGCCAGCGGGAGCGCTTCACCGCTTCACCGCCACCTAATCCGTACCTGAC 634  
QY 836 CACAGATGACACTGACACCCACCATCTCGCTGTCAGAGGGAGAGAGCCCACTAC 895  
Db 633 CACAGATGCA-CTGCGCCCGACCATCTCGCTGTCAGAGGG--AGAGCCCACTAC 577  
QY 896 CAGGCGCCCTGACCTTCACCTTCAGGAGCCCGAGAGCAGCAGCTGAACTGAACCGGAG 955  
Db 576 CAGGCGCCCTGACCTTCACCTTCAGGAGCCCGAGAGCAGCAGCTGAACTGAACCGGAG 517  
QY 956 TCGGTGGCGGACCCCAAGAGAACATCTTCAGACGTGACCTGATGATAGGCCAG 1015  
Db 516 TCGGTGGCGGACCCCAAGAGAACATCTTCAGACGTGACCTGATGATAGGCCAG 457  
QY 1016 CTGGGCGGCCCTGACCCCGCCAGCAGTAATCGGGCATCAGCGCCAGCTGACGGCAGC 1075  
Db 456 CTGGGCGGCCCTGACCCCGCCAGCAGTAATCGGGCATCAGCGCCAGCTGACGGCAGC 397  
QY 1076 GGGGGGCGATGAGAGGGGCGCCCGCCACCTACAGCAGAGGTATCGCCACTACCGGGG 1135  
Db 396 GGGGGGCGATGAGAGGGGCGCCCGCCACCTACAGCAGAGGTATCGCCACTACCGGGG 337  
QY 1136 TCGCTCTTCAGACCGACAGAGAGAGTGGGCGCCCTCTCTGAGGA--GGGAGCCCGGCT 1194  
Db 336 TCGCTCTTCAGACCGACAGAGAGAGTGGGCGCCCTCTCTGAGGA--GGGAGCCCGGCT 277  
QY 1195 CCACACACACATCGGCGCCCTAGAGAGCGACCATCTGAGAGCAAAAGAGAGATAA 1254  
Db 276 CCACACACACATCGGCGCCCTAGAGAGCGACCATCTGAGAGCAAAAGAGAGATAA 217  
QY 1255 ACAGAAAGAGACCTCTCTGAGGTTCCAGAGGGGGCCGGGCTGGGGCTGCTAGGTGAA 1314  
Db 216 ACAGAAAGAGACCTCTCTGAGGTTCCAGAGGGGGCCGGGCTGGGGCTGCTAGGTGAA 157  
QY 1315 AAGGAG 1321  
Db 156 AAGGAG 150  
RESULT 12



AF220208 651 bp mRNA linear ROD 05-MAR-2001  
LOCUS AF220208 Mus musculus Nedda WW domain-binding protein 4 mRNA, partial cds.  
DEFINITION AF220208  
ACCESSION AF220208.1 GI:12004973  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Jolliffe, C.N., Harvey, K.F., Haines, B.P., Parasiyam, G. and Kumar, S.  
TITLE Identification of multiple proteins expressed in murine embryos as  
binding partners for the WW domains of the ubiquitin-protein ligase  
Nedda  
JOURNAL Biochem. J. 351 Pt 3, 557-565 (2000)  
MEDLINE 20498735  
PUBMED 11042109  
REFERENCE  
AUTHORS Jolliffe, C.N. and Kumar, S.  
TITLE Direct Submission  
SUBMITTED (30-DEC-1999) Division of Haematology, Hanson Centre for  
Cancer Research, IMVS, Frome Rd, Adelaide, SA 5000, Australia  
LOCATION/Qualifiers  
FEATURES  
source  
1..651  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_lib="Novagen Cat. No. 69640-3"  
/dev\_stage="16 day embryo"  
<1..609  
/note="NAWBP4"  
/codon\_start=1  
/product="Nedda WW domain-binding protein 4"  
/protein\_id="FAG4247.1"  
/db\_xref="GI:12004974"  
/translation="SOARRRDLGSLSEGLTPESESTVSGMPEPVAYPRPTRLAY  
PPIQRREQPTPTQYALAPRITSLDGEPPYOGPCTLCRDEQLELRES  
VRAPPNRTIFDSDLIDSTMLGRCPPSSSGISATCYSSGSGMEGPPPTVSYGHYP  
GSSFOHQSSNPSSLLEGTRLHSHTAFLPNKKRKKOKRPL"  
BASE COUNT 146 a 234 c 176 g 95 t  
ORIGIN

Query Match 32.2%; Score 425; DB 10; Length 651;  
Best Local Similarity 83.2%; Pred. No. 2.4e-50;  
Matches 530; Conservative 0; Mismatches 80; Indels 27; Gaps 3;  
OY 641 AGCCAGGGGCGAGAGAGATGCTCTCAGAGAGATGCTGCGCCCTCGGAG 700  
1 AGCCAGGGGCGAGAGAGATGCTCTCAGAGAGATGCTGCGCCCTCGGAG 60  
OY 701 AGCAGGTGTCAAGCAAGGAATCCAGAGCGGAGGTCTAGCGCCCTCGGAG 760  
1 AGCAGGTGTCAAGCAAGGAATCCAGAGCGGAGGTCTAGCGCCCTCGGAG 117  
61 AGTACGGTGTCAAG--TGAATGCGGAGCGACAGGTCTAGCGCCCTCGGAG 117  
OY 761 GACCGCTGGCGGTGCGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGCC 820  
1 GACCGCTGGCGGTGCGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGCC 820  
DB 118 GACCGACTGCTGTGCGCCCTTCAGCGGAGCGGAGCGGAGCGGAGCGGAG 168  
OY 821 TATCCGTACTGAGAGAGATGCTGCTGAGAGAGATGCTGCGCCCTCGGAG 880  
1 TATCCGTACTGAGAGAGATGCTGCTGAGAGAGATGCTGCGCCCTCGGAG 228  
DB 169 TACCGTACTGAGAGAGATGCTGCTGAGAGAGATGCTGCGCCCTCGGAG 288  
OY 881 GAGCGCCCTCGGAGAGAGATGCTGCTGAGAGAGATGCTGCGCCCTCGGAG 940  
1 GAGCGCCCTCGGAGAGAGATGCTGCTGAGAGAGATGCTGCGCCCTCGGAG 288  
DB 229 GAGCGCCCTCGGAGAGAGATGCTGCTGAGAGAGATGCTGCGCCCTCGGAG 1000  
OY 941 GAACTGAACCGGAGTGTGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1000  
1 GAACTGAACCGGAGTGTGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAG 348  
DB 289 GAGTGAACCGGAGTGTGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1060  
OY 1001 ATGATAGTGTGAGGAGTGTGCGCGAGCGGAGCGGAGCGGAGCGGAGCGG 1060  
1 ATGATAGTGTGAGGAGTGTGCGCGAGCGGAGCGGAGCGGAGCGGAGCGG 408  
DB 349 ATAGACAGCAGCAGTGTGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAG 408

OY 1061 AGCTGTACAGGAGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 1120  
1 AGCTGTACAGGAGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 468  
DB 409 AGCTGTACAGGAGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 1180  
OY 1121 GAGCGCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 528  
1 GAGCGCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 528  
DB 469 GAGCGCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1240  
OY 1181 GAGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1240  
1 GAGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 573  
DB 529 GAGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1277  
OY 1241 AAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1277  
1 AAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610  
DB 574 AAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610

RESULT 13  
AX011709  
LOCUS AX011709 812 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 107 from Patent WO955858.  
ACCESSION AX011709  
VERSION AX011709.1 GI:9998233  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and  
Pillarsky, C.  
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue  
JOURNAL Patent: WO 995858-A 107 04-NOV-1999;  
SCHEMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GBS FUER GENOMFORSCHUN  
(DE); PILLARSKY CHRISTIAN (DE)  
LOCATION/Qualifiers  
FEATURES  
source  
1..812  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 157 a 272 c 237 g 146 t  
ORIGIN

Query Match 31.6%; Score 417.8; DB 6; Length 812;  
Best Local Similarity 77.4%; Pred. No. 2.3e-49;  
Matches 549; Conservative 0; Mismatches 122; Indels 38; Gaps 2;  
OY 251 GCGCGCGCGGAGATGCTGCGGCGGAGCGGAGCGGCGGCGGCGGCGG 310  
1 GCGCGCGCGGAGATGCTGCGGCGGAGCGGAGCGGCGGCGGCGGCGGCGG 189  
DB 130 GCTGTGAGCGGCTTCATCAGCGGCGGAGCGGAGCGGCGGAGCGGAG 370  
OY 311 GCGCGCGCTTCATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 249  
1 GCGCGCGCTTCATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 370  
DB 190 CTCAGAGAGATGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 430  
OY 371 CCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 309  
1 CCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 490  
DB 250 AGTCTTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 355  
OY 431 GTCAGAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 550  
1 GTCAGAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 415  
DB 310 -----CGTTTTCACCGCTTCAGAGCGGAGATGCTGCGAGCGGAG 610  
OY 491 AAGCGCTTTGTTCCAGAGAGATGAGATGAGAGAGATGAGAGATGAGATG 610  
1 AAGCGCTTTGTTCCAGAGAGATGAGATGAGAGAGATGAGAGATGAGATG 415  
DB 356 AAGCGCTTTGTTCCAGAGAGATGAGATGAGAGAGATGAGAGATGAGATG 610  
OY 551 ATCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 451  
1 ATCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 610  
DB 416 ATCGTGTGTG-----TACGTTGCTGCTGAGCGGAGCGGAGCGGAG 670  
OY 611 CTGCTGTGAGGAGTCTTCATCAGCGGCGGAGCGGAGCGGAGCGGAGCGG 670



```

Db 452 CTGCTCTCAGCGTCTCTCATACAGCCGACAGCGAGGGGCGAGAGAGAAGATGCCCTG 511
Qy 671 TCTCTGAGAAGATGCTCTGTGGCCCTCGAGAGACACATGTCAGCAACGGAATCCGAG 730
Db 512 TCTCTGAGAAGATGCTCTGTGGCCCTCGAGAGACACATGTCAGCAACGGAATCCGAG 571
Qy 731 CCGAGAGTCTACGCGCCGCTCGAGCCGACGACGCGCTGCGCGCCGCTTGGCCAG 790
Db 572 CCGAGAGTCTACGCGCCGCTCGAGCCGACGACGCGCTGCGCGCCGCTTGGCCAG 631
Qy 791 CCGAGAGCTTCCACCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTT 850
Db 632 CCGAGAGCTTCCACCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTT 691
Qy 851 CCAACCCACCTCTCTGCTGTAGAGCGGGAGAGAGCCGCTTCCAGCCGCTTCCAGCCGCTT 910
Db 692 CCGAGAGCTTCTCTGCTGTGTAGAGCGGGAGAGAGCCGCTTCCAGCCGCTTCCAGCCGCTT 751
Qy 911 CTCGAGCTTGGGAGCCGCGAGCAGCAGCTGGAACCTGAACCGGGAGTGG 959
Db 752 TTCAAGGTTGGGAGCCCGAGAGAGAGTGTGAATTGAACGGGGATTGG 800

RESULT 14
AX071267 408 bp DNA linear PAT 25-JAN-2001
LOCUS Sequence 1739 from Patent WO0102568.
DEFINITION AX071267
ACCESSION AX071267
VERSION AX071267.1 GI:12581618
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 408)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klingner,J.,
Lasson,G., Drmanac,R., Randazzo,C., Kennedy,G.C., Pot,D.,
Labat,I., Leshkovitz,D., Kila,D., Garcia,V. and Strache-Grain,B.
HUMAN genes and gene expression products
Patent: WO 0102568-A 1739 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES
source
1..408
location/Qualifiers
BASE COUNT 81 a 159 c 115 g 53 t
RIGIN
Query Match 30.4%; Score 401.4; DB 6; Length 408;
Best Local Similarity 99.8%; Pred. No. 5.2e-47;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 581 GTGATCAGTGGCTGTGAGCACTACAGCTGTCTCAGCGTCTCTCATCAGCGGCGAC 640
Db 6 GAGATCAGTGGCTGTGAGCACTACAGCTGTCTCAGCGTCTCTCATCAGCGGCGAC 65
Qy 641 AGCCAGGGGGGAGAGAGAGATGCTCTCTCAGAGAGATGCTGTGGCCCTCGAG 700
Db 66 AGCCAGGGGGGAGAGAGAGATGCTCTCTCAGAGAGATGCTGTGGCCCTCGAG 125
Qy 701 AGCAGAGTGTGAGCAAGGATCCGAGCGGAGGTCTACAGCCGCGCGCCGCGCCAGC 760
Db 126 AGCAGAGTGTGAGCAAGGATCCGAGCGGAGGTCTACAGCCGCGCGCCGCGCCAGC 185
Qy 761 GAGCGCTGCGCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTT 820
Db 186 GAGCGCTGCGCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTTCCAGCCGCTT 245
Qy 821 TATCCGTAAGTGTGAGCAAGGATGAGCTGCAACCCACCATCTGTGCTCAGAGGGAG 880
Db 246 TATCCGTAAGTGTGAGCAAGGATGAGCTGCAACCCACCATCTGTGCTCAGAGGGAG 305

```

```

Qy 881 GAGCCCCCACCCTACAGAGGCGCCCTGCACCTCAGCTTGGGAGCCCGAGCAGACTG 940
Db 306 GAGCCCCCACCCTACAGAGGCGCCCTGCACCTCAGCTTGGGAGCCCGAGCAGACTG 365
Qy 941 GAACTGAACCGGAGTGTGTCGCGCACCCCAACCAAGAACCA 983
Db 366 GAACTGAACCGGAGTGTGTCGCGCACCCCAACCAAGAACCA 408

RESULT 15
AL837509 156075 bp DNA linear HTG 17-AUG-2002
LOCUS Mus musculus chromosome 2 clone RP23-44L6, *** SEQUENCING IN
DEFINITION PROGRESS ***, 4 unordered pieces.
ACCESSION AL837509
VERSION AL837509.4 GI:22416197
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 156075)
AUTHORS Almeida,J.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Aug 21, 2002 this sequence version replaced gi:22204793.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: bm44l6
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Consensus quality: 155649 bases at least Q40
Consensus quality: 155716 bases at least Q30
Consensus quality: 155743 bases at least Q20
Insert size: 155775; sum-of-coverage
Insert size: 162242; 2.6% error; agarose-fp
Quality coverage: 7.64x in Q20 bases; sum-of-coverage
Quality coverage: 7.40x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 9981: contig of 9981 bp in length
* 9982 10081: gap of 100 bp
* 10082 13090: contig of 3009 bp in length
* 13091 13190: gap of 100 bp
* 13191 95958: contig of 82768 bp in length
* 95959 96058: gap of 100 bp
* 96059 156075: contig of 60017 bp in length.
location/Qualifiers
1..156075
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-44L6"
/clone_1lb="RPCT-23"
1..9981
/note="assembly_fragment:00303
fragment_chain:1"
10082..13090
misc_feature
misc_feature
misc_feature

```



	45	72.2	5.5	2307	24	ABK32842	DNA encoding human
						ALIGNMENTS	
						RESULT 1	
						ABK12137	
						ID ABK12137 standard; cDNA; 1321 BP.	
						AC ABK12137;	
						DT 05-JUN-2002 (first entry)	
						XX Human cDNA encoding Mechanically Induced Vascular Receptor 1, MIVR-1.	
						DE Human; ss; gene; MIVR-1; Mechanically Induced Vascular Receptor 1;	
						KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;	
						KM cardiac cell; anti-apoptotic; vascular endothelial cell;	
						KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;	
						KM heart failure.	
						OS Homo sapiens.	
						XX Location/Qualifiers	
						FH CDS 413..1276	
						FT /*tag= a	
						FT /product= "MIVR-1"	
						/note="This region is specifically claimed in claim 3"	
						WO200216416-A2.	
						PN 28-FEB-2002.	
						PD 21-AUG-2001: 2001MO-US26089.	
						PF 22-AUG-2000: 2000US-227159P.	
						PR XX	







DR P-PSDB; AAB18461.  
 XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX  
 XX Disclosure; Page -: 175pp; English.  
 PS  
 CC AAA75163-65 encode human TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC modulate the proliferation, differentiation, and/or function of cells  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, atelectasis,  
 CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral oedema,  
 CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.  
 CC  
 CC  
 XX  
 SO Sequence 969 BP: 211 A; 317 C; 293 G; 148 T; 0 other;

Query Match 60.7%; Score 801.6; DB 21; Length 969;  
 Best Local Similarity 99.5%; Pred. No. 2.8e-120;  
 Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 514 GGAATATCAGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTATGATGATGAT 573  
 DB 2 GGAATATCAGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTATGATGATGAT 61  
 QY 574 GGTGGTGTGATCAGTGCCTGCTGAGCACTACAGTGTCTCAGAGGTCCTCATCAG 633  
 DB 62 GGTGGTGTGATCAGTGCCTGCTGAGCACTACAGTGTCTCAGAGGTCCTCATCAG 121  
 QY 634 CCGGACACAGCCAGGGGGGAGAGAGATGCCCTGTCTCAGAGAGATGCTGTGGCC 693  
 DB 122 CCGGACACAGCCAGGGGGGAGAGAGATGCCCTGTCTCAGAGAGATGCTGTGGCC 181  
 QY 694 CTGGAGAGCAGATGTCAGGCAAGCAATCCAGAGCCGAGGTCTACGCCGCCCTCG 753  
 DB 182 CTGGAGAGCAGATGTCAGGCAAGCAATCCAGAGCCGAGGTCTACGCCGCCCTCG 241  
 QY 754 GCCACAGCAGCCGCTGGCGCTGCGCCCTTGCCAGGGAGGAGCTTCCACGCCCTTCCA 813  
 DB 242 GCCACAGCAGCCGCTGGCGCTGCGCCCTTGCCAGGGAGGAGCTTCCACGCCCTTCCA 301  
 QY 814 GCCACCTATCCGATCTGACAGCAGAGATGACCTGCAACCCACATCTTGCTGTGAGA 873  
 DB 302 GCCACCTATCCGATCTGACAGCAGAGATGACCTGCAACCCACATCTTGCTGTGAGA 361  
 QY 874 CGGGAGAGAGCCGCCACCTTACAGGGCCCTGCAACCTTGGAGAGCCCGAGAGA 933  
 DB 362 CGGGAGAGAGCCGCCACCTTACAGGGCCCTGCAACCTTGGAGAGCCCGAGAGA 421  
 QY 934 GCACCTGGAATGAACCGGAGTGGTGGCGGACCCCAAGCAAGCAATCTTGCAGAG 993  
 DB 422 GCACCTGGAATGAACCGGAGTGGTGGCGGACCCCAAGCAAGCAATCTTGCAGAG 481  
 QY 994 TGACCTGATGATAGTGCACAGCTGGGGGCGCCCTGCCCCCCAGTAATCTCGGGCAT 1053  
 DB 1053 TGACCTGATGATAGTGCACAGCTGGGGGCGCCCTGCCCCCCAGTAATCTCGGGCAT 1053

DB 482 TGACCTGATGATAGTGCACAGCTGGGGGCGCCCTGCCCCCCAGTAATCTCGGGCAT 541  
 QY 1054 CAGCGCCACGAGTGTACAGGAGCGGCGGCGCATGAGAGGGCGCGGCCACCTACAGCGA 1113  
 DB 542 CAGCGCCACGAGTGTACAGGAGCGGCGGCGCATGAGAGGGCGCGGCCACCTACAGCGA 601  
 QY 1114 GGTATATGCGGCACATACCCGGGGTCTCTTCCACACAGCAGAGAGAGTGGCGCCCTC 1173  
 DB 602 GGTATATGCGGCACATACCCGGGGTCTCTTCCACACAGCAGAGAGAGTGGCGCCCTC 661  
 QY 1174 CTTCGTGAGGGAGCCCGGCTCCACACACACATGCGCCCTTACAGAGCGAGCCAT 1233  
 DB 662 CTTCGTGAGGGAGCCCGGCTCCACACACACATGCGCCCTTACAGAGCGAGCCAT 721  
 QY 1234 CTGGAGCAAGAGAGATTAACAGAAAGACACCTCTTAGGGTCCCAAGGGGGCGG 1293  
 DB 722 CTGGAGCAAGAGAGATTAACAGAAAGACACCTCTTAGGGTCCCAAGGGGGCGG 781  
 QY 1294 GGCTGGGGCTGCTAGGTGAAGAGCAG 1321  
 DB 782 GGCTGGGGCTGCTAGGTGAAGAGCAG 809  
 RESULT 5  
 ID AAA75164 standard; cDNA; 969 BP.  
 AC AAA75164;  
 XX 15-JAN-2001 (first entry)  
 DE cDNA clone encoding a human TANGO 261 polypeptide.  
 KW TANGO 266; TANGO 216; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 6..764  
 FT /tag= a  
 FT /product= "TANGO 261"  
 PN WO200052022-A1.  
 XX PD 08-SEP-2000.  
 PF 01-MAR-2000; 2000WO-US05226.  
 XX PR 01-MAR-1999; 99US-0122458.  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 DR WPI; 2000-579269/54.  
 DR P-PSDB; AAB18462.  
 XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX Disclosure; Page -: 175pp; English.  
 XX AAA75163-65 encode human TANGO 261 proteins. The specification also



describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO polypeptides can be used to modulate cellular proliferation, modulate cellular differentiation and/or modulate cellular adhesion. The proteins can be used to treat any von Willebrand factor-associated disorder, regulate extracellular matrix structuring, cellular adhesion, and cell trafficking and/or migration, modulate cellular interactions, modulate cell adhesion in proliferative disorders, such as cancer, modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, and leukocytes, treat bone marrow, blood and hematopoietic associated diseases and disorders, atelectasis, bronchial pulmonary congestion or oedema, emphysema, chronic bronchitis, asthma and bronchiectasis, intestinal disorders, spleen associated diseases, modulate renal disorders, treat cardiovascular disorders such as ischemic heart disease, modulate the proliferation, differentiation, and/or function of bone and cartilage cells and to treat bone and/or cartilage associated diseases or disorder. They may also be used to treat disorders associated with the ovaries, and cerebral oedema, hydrocephalus, brain herniations, iatrogenic disease, inflammations, bacterial and viral meningitis, Alzheimer's disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers, hydrocephalus and encephalitis, and treat hepatic disorders.

note: the present sequence does not appear in the specification; it was created using information provided.

Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;

Query Match 60.7%; Score 801.6; DB 21; Length 969;  
Best Local Similarity 99.5%; Pred. NO. 2.8e-120;  
Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 514 GGAGATCAGCAGCTGAGTGTGTCAGATCATCATCTGCTGTGTGATGATGATGAT 573
DB 2 GGAGATGCGCGAGCTGAGTGTGTCAGATCATCATCTGCTGTGTGATGATGATGAT 61
OY 574 GGTGTGTGTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633
DB 62 GGTGTGTGTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
OY 634 CCGGACAGCAGCGGCGGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693
DB 122 CCGGACAGCAGCGGCGGAGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
OY 694 CTGCGAGAGCAGAGTGTGAGGACGATCCAGAGCGGAGGTCTAGCGCCGCTGCG 753
DB 182 CTGCGAGAGCAGAGTGTGAGGACGATCCAGAGCGGAGGTCTAGCGCCGCTGCG 241
OY 754 GCCCAGCAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 813
DB 242 GCCCAGCAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 301
OY 814 GCCCAGCTATCGTACCTGCGAGCAGAGATGACCTGCGACCCACCATCTGCGTGCAG 873
DB 302 GCCCAGCTATCGTACCTGCGAGCAGAGATGACCTGCGACCCACCATCTGCGTGCAG 361
OY 874 CGGGAGAGAGCGCCGACCTACAGAGGCGCTGCGACCTGCGAGCTTTCGGAGCCCGAGCA 933
DB 362 CGGGAGAGAGCGCCGACCTACAGAGGCGCTGCGACCTGCGAGCTTTCGGAGCCCGAGCA 421
OY 934 GCAGCTGGAAGTGAACCGGAGTGTGCTGCGGACCCGCAACAGACATCTTTCAGAG 993
DB 422 GCAGCTGGAAGTGAACCGGAGTGTGCTGCGGACCCGCAACAGACATCTTTCAGAG 481
OY 994 TGACCTGATGATGATGTCAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAT 1053
DB 482 TGACCTGATGATGATGTCAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAT 541
OY 1054 CAGCGCCAGCTGCTACGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAT 1113
DB 542 CAGCGCCAGCTGCTACGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAT 601
OY 1114 GGTCTATGCGCAGCTACCGCGGCGGCTCTCTTCAGCAGCAGCAGAGCAGTGGGCGGCGCTC 1173
DB 602 GGTCTATGCGCAGCTACCGCGGCGGCTCTCTTCAGCAGCAGCAGAGCAGTGGGCGGCGCTC 661

```

```

OY 1174 CTGCTGAGAGGAGACCCGCTCCACACACACATGCGGCGGCGGCGGCGGCGGCGGAT 1233
DB 662 CTGCTGAGAGGAGACCCGCTCCACACACACATGCGGCGGCGGCGGCGGCGGCGGAT 721
OY 1234 CTGAGCAGAAAGAGAGTAAACAGAAAGAGACACCTCTTAGGCTCCCGAGGGGCGG 1293
DB 722 CTGAGCAGAAAGAGAGTAAACAGAAAGAGACACCTCTTAGGCTCCCGAGGGGCGG 781
OY 1294 GCGTGGGCGCTCGTAGCTGAGTGAAGAGCAG 1321
DB 782 GCGTGGGCGCTCGTAGCTGAGTGAAGAGCAG 809

RESULT 6
ID AAA75165 standard; cDNA: 969 BP.
AC AAA75165;
XX 15-JAN-2001 (first entry)
DE cDNA clone encoding a human TANGO 261 polypeptide.
XX
XX TANGO 266; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cell trafficking; cancer;
KW von Willebrand factor-associated disorder; atelectasis; pulmonary congestion;
KW hematopoietic associated disease; chronic bronchitis; bronchial asthma; bronchiectasis;
KW oedema; emphysema; spleen associated disease; renal disorder;
KW intestinal disorder; ischemic heart disease; hydrocephalus;
KW cardiovascular disorder; iatrogenic disease; inflammation; meningitis;
KW brain herniation; Parkinson's disease; cerebral toxoplasmosis; Alzheimer's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 6..764
FT CDS /tag= a
FT /product= "TANGO 261"
FT
XX
XX W0200052022-A1.
XX
XX 08-SEP-2000.
XX
XX 01-MAR-2000; 2000WO-US05226.
XX
XX 01-MAR-1999; 99US-0122458.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX
XX WPI: 2000-579269/54.
XX P-PSDB: AAB18463.
XX
XX Novel human and murine secreted proteins designated TANGO 216, 261,
XX 262, 266 and 267 useful as modulating agents of cellular processes,
XX e.g. for treating cancer -
XX
XX Disclosure; Page -, 175pp; English.
XX
XX AAA75163-65 encode human TANGO 261 proteins. The specification also
XX describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
XX polypeptides can be used to modulate cellular proliferation, modulate
XX cellular differentiation and/or modulate cellular adhesion. The
XX proteins can be used to treat any von Willebrand factor-associated
XX disorder, regulate extracellular matrix structuring, cellular adhesion,
XX and cell trafficking and/or migration, modulate cellular interactions,
XX modulate cell adhesion in proliferative disorders, such as cancer,
XX modulate the proliferation, differentiation, and/or function of cells
XX that appear in the bone marrow, and leukocytes, treat bone marrow, blood

```



CC and hemopoietic associated diseases and disorders, atelectasis,  
CC pulmonary congestion or edema, emphysema, chronic bronchitis, bronchial  
CC asthma and bronchiectasis, intestinal disorders, spleen associated  
CC diseases, modulate renal disorders, treat cardiovascular disorders such  
CC as ischemic heart disease, modulate the proliferation, differentiation,  
CC and/or function of bone and cartilage cells and to treat bone and/or  
CC cartilage associated diseases or disorder. They may also be used to  
CC treat disorders associated with the ovaries, and cerebral oedema,  
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
CC bacterial and viral meningitis, Alzheimer's disease, cerebral  
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
CC hydrocephalus and encephalitis, and treat hepatic disorders.  
CC note: the present sequence does not appear in the specification; it was  
CC created using information provided.

50 Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other,

Query Match	Score	DB	Length
60.78;	801.6;	DB 21;	969;

Best Local Similarity 99.5%; Pred. No. 2.8e-120;

Matches 804; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY	514	GGAGATACGGAGAGCTGGAGATTGTTGTCAGATCAATCAATCACTGCTGATATGATGAT	573
Dp	2	GGAAATGGCGAGGCTGGAGATTGTTGTCAGATCAATCAATCACTGCTGATATGATGAT	61
OY	574	GGTGGTGGTATCATCTGCTGCTGCTGAGACCCTCAAGCTGTCTGCAAGGCTCTTCATCAG	633
Dp	62	GGTGGTGGTATCATCTGCTGCTGCTGAGACCCTCAAGCTGTCTGCAAGGCTCTTCATCAG	121
OY	634	CCGGCACAGCCAGGGGCGGAGAGAGAAATGCCCTGTCTCTAGAAAGATGCCCTGTGGCC	693
Dp	122	CCGGCACAGCCAGGGGCGGAGAGAGAAATGCCCTGTCTCTCTAGAGAGGATGCCCTGTGGCC	181
OY	694	CTCGAGAGCACTGTGTACAGCAACGGAATCCAGAGCCGAGGTGTACGCCGCCGCTCG	753
Dp	182	CTCGAGAGCACTGTGTACAGCAACGGAATCCAGAGCCGAGGTGTACGCCGCCGCTCG	241
OY	754	GCCACACGACGCGCTGTGGCCGCTGTGCCGCTTGTGCCACAGCGGAGCGCTTCACCGCTTCCA	813
Dp	242	GCCACACGACGCGCTGTGGCCGCTGTGGCCGCTTGTGCCACAGCGGAGCGCTTCACCGCTTCCA	301
OY	814	GCCCACTATCCGTACTGTGCAGCAGCAGATGCACCTGTCCACCCACCATCTCGCTGTGAGA	873
Dp	302	GCCCACTATCCGTACTGTGCAGCAGCAGATGCACCTGTCCACCCACCATCTCGCTGTGAGA	361
OY	874	CGGGGAGGAGGCCCCCAGCCCTTACACAGGGGCCCCCTGACACCTTCACCTTCGGGAGCCCGAGAGA	933
Dp	362	CGGGGAGGAGGCCCCCAGCCCTTACACAGGGGCCCCCTGACACCTTCACCTTCGGGAGCCCGAGAGA	421
OY	934	GCACCTGGAACTGAACCGGGGAGTGTGTGCGCGACGCCCAAAGAGAAACCATCTTGCAGAG	993
Dp	422	GCACCTGGAACTGAACCGGGGAGTGTGTGCGCGACGCCCAAAGAGAAACCATCTTGCAGAG	481
OY	994	TGACCTGTATGATATGTGCCAAGGCTGGGGGCGCCCTGTGCCCCCGACAGCATCTCGGGCAT	1053
Dp	482	TGACCTGTATGATATGTGCCAAGGCTGGGGGCGCCCTGTGCCCCCGACAGCATCTCGGGCAT	541
OY	1054	CAGGGCCACGTGTACGGGCGAGCGGGGGGGGCGCATGTGAGGGGGCGGCCGCCCACTACAGCGA	1113
Dp	542	CAGGGCCACGTGTGTACGGGCGAGCGGGGGGGGCGCATGTGAGGGGGCGGCCGCCCACTACAGCGA	601
OY	1114	GGTCATCGGCACTAACCCCGGGGGTCTCTCTTCACAGACACAGCAGAGCATGTGGGGCGGCGCTC	1177
Dp	602	GGTCATCGGCACTAACCCCGGGGGTCTCTCTTCACAGCAGCAGAGCATGTGGGGCGGCGCTC	661
OY	1174	CTTGCTGGAGGGGAGCCGGGCTCCACACACACACATGTGCGCCCTTAGAGAGCGCAGCCAT	1233
Dp	662	CTTGCTGGAGGGGAGCCGGGCTCCACACACACATGTGCGCCCTTAGAGAGCGCAGCCAT	721
OY	1234	CTTGAGGCAAAAGAAAGGATTAACAGAAAGACACCTCTTATAGGGTCCCAAGGGGGGCGG	1293
Dp	722	CTTGAGGCAAAAGAAAGGATTAACAGAAAGACACCTCTTATAGGGTCCCAAGGGGGGCGG	781

QY 1294 GGCTGGGGCTGCGTAGGTGAAGAGCAG 1321  
|||  
Db 782 GGCTGGGGCTGCGTAGGTGAAGAGCAG 809

RESULT 7	
AAA47429	
ID	AAA47429 standard; DNA; 1061 BP
XX	
AC	AAA47429;
XX	
DT	20-OCT-2000 (first entry)

DE Sequence encoding human neuron-associated protein

KW Neuromuscular protein; NEAP; neurological disorder; epilepsy;  
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
KW Alzheimer's disease; Pick's disease; Huntington's disease;  
KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
KW prion disease; Kurru; Creutzfeldt-Jakob disease; neurofibromatosis;  
KW cerebral palsy; muscular dystrophy; central nervous system;  
KW peripheral nervous system; PNS; myopathy; schizophrénia;  
KW acral keratosis; arteriosclerosis; atherosclerosis; MCTD;  
KW cirrhosis; hepatitis; mixed connective tissue disease; BCRT;  
KW myofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KW AIDS; Addison's disease; adult respiratory distress syndrome;  
KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KW Werner syndrome; trauma; human; ds.

OS Homo sapiens.

	Key	Location/Qualifiers
FH	CDS	101..859
FT		

PN WO200034477-A2

PD 15-JUN-2000

PF 10-DEC-1999; 99WO-US30408.

11-DEC-1998; 9803-0210083  
PR 11-DEC-1998; 9803-9123456

PR	03-FEB-1999;	99US-0119303
PR	16-MAR-1999;	99US-0124687

AA  
PA (INCY-) INCYTE PHARM INC.

aa Tang YT, Yue H, Baughn MR,  
PI

PL LU DAM, AZIMIZDA I;  
XX

DR WFL; 2000-423423/30.  
DR P-PSDB: AAB01388.

PT      New human neuron-associated

disorders including cancer, PT

PS Claim 9; Page 136; 145pp; Er

Human neuron-associated prot

expression or activity of NH

treating or preventing disorder associated with increased expression or activity of NEAP. NEAP or their fragments or derivatives are useful for treating neurological disorders such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEAPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jacob disease, nutritional and

Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 other;

0

Qy	1219	AGAGACCGGAGCGC	CTGTGAGCAAGAGAGATTAACAGAAAGACACCTCTCTAGGG	1278
Db	802	AGAGAGCGCAGCCATCT	TGTGAGCAAGAGAGATTAACAGAAAGACACCTCTCTAGGG	861
Qy	1279	TCCCCAGGGGGCGCGGCTG	GGGCTGCCGTAGGTGAAAAAGGCAG	1321
Db	862	TCCCCAGGGGGCGCGGCTG	GGGCTGCCGTAGGTGAAAAAGGCAG	904
RESULT 8				
ID	AA157868	AA157868 standard; cDNA; 1066 BP.		
AC	AA157868;			
XX				
DT	22-OCT-2001	(first entry)		
XX				
DE	Human polynucleotide SEQ ID NO 71.			
XX				
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;			
KW	peripheral nervous system; neuropathy; central nervous system; CNS;			
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;			
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;			
KW	chemokineic; thrombolytic; drug screening; arthritis; inflammation;			
KW	leukaemia; ss.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200153312-A1.			
XX				
PD	26-JUL-2001.			
XX				
PF	26-DEC-2000; 2000WO-US34263.			
XX				
PR	21-JAN-2000; 2000US-0488725.			
PR	25-APR-2000; 2000US-0552317.			
PR	09-JUL-2000; 2000US-0598042.			
PR	19-JUL-2000; 2000US-0620312.			
PR	03-AUG-2000; 2000US-0653450.			
PR	14-SEP-2000; 2000US-0662191.			
PR	19-OCT-2000; 2000US-0693036.			
PR	29-NOV-2000; 2000US-0727344.			
PA	(HYSE-) HYSEQ INC.			
XX				
PL	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;			
PL	Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;			
PL	Zhao QA, Zhou P, Goodrich R, Drmanac RT;			
XX				
DR	WPI: 2001-442253/47.			
DR	P-PSDB; AAM38712.			
PT	Novel nucleic acids and polypeptides, useful for treating disorders			
PT	such as central nervous system injuries -			
XX				
PS	Claim 1; SEQ ID NO 71; 10078bp; English.			
CC				
CC	The invention relates to human nucleic acids (AA157798-AA161369) and			
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,			
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful			
CC	in gene therapy. A composition containing a polypeptide or polynucleotide			
CC	of the invention may be used to treat diseases of the peripheral nervous			
CC	system, such as peripheral nervous injuries, peripheral neuropathy and			
CC	localised neuropathies and central nervous system diseases, such as			
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic			
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the			
CC	utilisation of the activities such as: immune system suppression,			
CC	Activin/Inhibin activity, chemotactic/chemokineic activity, haemostatic			
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,			
CC	assays for receptor activity, arthritis and inflammation, leukemias and			
CC	C.N.S disorders.			
CC	Note: The sequence data for this patent did not form part of the printed			

Note: The sequence data for this patent did not form part of the printed













... diseases, modulate renal disorders, treat and





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 12:18:54 ; Search time 63.5678 Seconds  
(without alignments)  
6373.035 Million cell updates/sec

Title: US-09-934-249-1

Perfect score: 1321

Sequence: 1 cgaccgcgctccgagcgca.....ctgcgtagtgtgaagcgag 1321

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Optical number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCBUS.COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfilesl.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352.2	26.7	921	4	US-09-091-952A-7
2	352.2	26.7	8065	4	US-09-091-952A-6
3	284.2	21.5	867	4	US-09-091-952A-8
4	68	5.1	44377	2	US-08-804-227C-7
5	68	5.1	44377	2	US-08-804-198-1
6	67.6	5.1	4257	2	US-08-690-473-1
7	67.6	5.1	4257	4	US-09-259-821A-1
8	67.6	5.1	4257	4	US-08-843-659-1
9	67.6	5.1	12001	1	US-08-458-568A-11
10	66	5.0	152331	3	US-09-128-155-16
11	65.4	5.0	4403765	4	US-09-103-840A-2
12	64	4.8	4524	2	US-08-845-998-7
13	64	4.8	4524	3	US-09-206-537-7
14	64	4.8	4524	4	US-09-430-854-7
15	63.2	4.8	2991	3	US-08-795-430-48
16	63.2	4.8	2991	4	US-09-355-700-48
17	62.2	4.7	319	4	US-09-165-264-8
18	62.2	4.7	8438	1	US-07-945-283-1
19	62	4.7	8438	1	US-07-945-283-1
20	61.4	4.6	4060	1	US-08-308-949A-1
21	61	4.6	4411529	4	US-09-103-840A-1
22	60.4	4.6	320	4	US-09-165-264-13
23	59.8	4.5	320	4	US-09-165-264-14
24	59.4	4.5	8147	4	US-09-514-247A-9
25	59.2	4.5	320	4	US-09-165-264-7
26	59.2	4.5	71989	4	US-09-443-501A-2
27	59	4.5	801	2	US-08-770-379-16

28	59	4.5	801	4	US-08-757-669A-16	Sequence 16, Appl
29	59	4.5	801	4	US-09-298-568-3	Sequence 3, Appl
30	59	4.5	801	4	US-09-230-371A-16	Sequence 16, Appl
31	58.4	4.4	320	4	US-09-165-264-11	Sequence 11, Appl
32	58.2	4.4	15872	4	US-09-105-537-1	Sequence 1, Appl
33	58.2	4.4	43280	2	US-08-804-227C-1	Sequence 1, Appl
34	58	4.4	44377	2	US-08-804-227C-7	Sequence 7, Appl
35	58	4.4	44377	2	US-08-804-198-1	Sequence 1, Appl
36	57.6	4.4	318	4	US-09-165-264-12	Sequence 12, Appl
37	56.8	4.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl
38	56.4	4.3	1877	4	US-09-780-173A-10	Sequence 10, Appl
39	56.2	4.3	1209	6	5352575-4	Patent No. 5352575
40	56.2	4.3	1213	4	US-09-232-468A-7	Sequence 7, Appl
41	56	4.2	4488	4	US-08-406-030A-3	Sequence 3, Appl
42	56	4.2	5434	2	US-08-841-349-1	Sequence 1, Appl
43	56	4.2	12001	1	US-08-458-568A-11	Sequence 11, Appl
44	55.6	4.2	2483	1	US-08-464-340A-3	Sequence 3, Appl
45	55.6	4.2	2483	5	PCT-US94-08449A-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-09-091-952A-7  
Sequence 7, Application US/09091952A  
Patent No. 6458532  
GENERAL INFORMATION:  
APPLICANT: Detera-Wadleigh, Sevilla D.  
Gershon, Elliot S.  
Badner, Judith A.  
Goldin, Lynn R.  
Berrettini, Wade H.  
Yoshikawa, Takeo  
Sanders, Alan R.  
Esterling, Lisa E.  
TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,952A  
FILING DATE: 19-Apr-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,278  
FILING DATE: 28-Oct-1996  
APPLICATION NUMBER: PCT/US97/19381  
FILING DATE: 28-Oct-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 015280-29710005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

Query Match	26.7%	Score 352.2	DB 4	Length 921
Best Local Similarity	73.1%	Pred: No. 5.4e-58		
Matches 482	Conservative 0	Mismatches 168	Indels 9	Gaps 2
TOPOLOGY: linear MOLECULE TYPE: cdna FEATURE: NAME/KEY: CDS LOCATION: 1...921 OTHER INFORMATION: Clone 22 coding region SEQUENCE DESCRIPTION: SEQ ID NO: 7: us-09-091-952A-7				
OY	506	CAGACATGGAATCAGGAGCTGGAGTGTGTCAGATCATCATCATGCTGGTGGATG	565	
DB	166	CCGGGCACTTCACTCAGAGCTGGAGTTCGCCCAATCATATCATCTCGGGGGTTC	225	
OY	566	ATGTAATGTTGGTGGTGTGATACAGTCCCTGCTAGACCACTACAAAGCTGTCTCAGCGTCC	625	
DB	226	ACGGGATGATGGTGGTGTGATCTGTGCTGCTGTAACCACTACAAAGTCTCCACGGGCTCC	285	
OY	626	TTTCATCACCGGCAACGCGGGGGGAGAGAGAGAAATGATGCCCTGTCTCTCAAGAGATCC	685	
DB	286	TTTCATCAACCGGCGCCCAACCGAGAGCGGAGAGCGAGAGAGAGGAGGCTCCGAGAAAGGCTCC	345	
OY	686	CTGTGCGCCCTCGAGAGAGACAGTGTTCAGAGCAAGGAAATCCAGAGCGGAGGCTCTACGCC	745	
DB	346	CTGTGCGCTTCAGACAGGCGCGCACCGGGGGCTGGG-----GGCTGGAGATCATGCAAT	399	
OY	746	CCGCGCTCGGCCACCGAGCCGCTGGCGCTGCCCTTCGCCCGCCAGGGAGCGCTTCCAC	805	
DB	400	GCCCGCGGCTCCAGGAGACAGGTTCAACAGCGCGCTCTTCATCAAGAGGATCGCTTCACG	459	
OY	806	CGCTTCCAGCCCACTATCCGATCCGATCCGACAGACAGATGACCTGCGACCCACCATCTCC	865	
DB	460	CGCTTCCAGCCCACTATCCGATCCGATCCGACAGACAGATGATGATCTTCTCCACCATCTCC	519	
OY	866	CTGTGACAGCGGGGAGAGAGCCCGCCCTTACACAGGCGCCCTGACCCCTCCAGCTTGGGAGC	925	
DB	520	CTGTCCGAGCGGGAAGAGACACCTCTTACACAGGGGGCTCTGCACCCCTGACGCTCCGGGAGC	579	
OY	926	CCCGAGCAGCACTGGAAGTGAACCGGGGGTGGGCGCGGCGACCCGCCAAGCAAGCAACCAAC	985	
DB	580	CCTGAAACAGCAATGGAATCACTCAACCGAGAGATCTGTGAGGGGCCCAACCCCAACGGAACATA	639	
OY	986	TTTCGACAGTGAACCTGATGATAGTCCAGG---CTGGGGCGGCCCTTGCCCCCCCCAGCACT	1042	
DB	640	TTTGCACATGATTTTAATAGACATTTCTATGTATATGCGGGGGTTCATGCCCCACCCAGCAACG	699	
OY	1043	AACCTGGGCACTCAGGCGGCAACGTCATACGCGGAGCGGCGGCGCATGGAAGGGGGCGCGCC	1102	
DB	700	AACCTGGGCACTCAGTGCACAGCACTCGCAGCAATTAACGGGAGGATGAGGGGGCGCACCCCC	759	
OY	1103	ACCTACAGCGAGGTCATCGGCACTACCCCGGGGTCCTCTCTTCACACACGAGAGAGAG 1161		
DB	760	ACATACAGCGAGGTCATGAGGCGCACCAACCGAGCGGCTCTTCTCTCATCACACAGGCGAG 818		
RESULT 2 us-09-091-952A-6 Sequence 6, Application us/09091952A Patent No. 6458532 GENERAL INFORMATION: APPLICANT: Detera-Wadleigh, Sevilla D. Cersnon, Elliot S. Badner, Judith A. Goldin, Lynn R. Berretini, Wade H. Yoshikawa, Takeo Sanders, Alan R. Esterling, Lisa E. TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness				

```

? NUMBER OF SEQUENCES: 197
? CORRESPONDENCE ADDRESS:
? ADDRESS: Townsend and Townsend and Crew LLP
? STREET : Two Embarcadero Center, Eighth Floor
? CITY : San Francisco
? STATE : CA
? COUNTRY: USA
? ZIP: 94111-3834
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq for Windows Version 2.0
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/091,952A
? FILING DATE: 19-Apr-1999
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/029,278
? FILING DATE: 28-OCT-1996
? APPLICATION NUMBER: PCT/US97/19381
? FILING DATE: 28-OCT-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Smith, Timothy L.
? REGISTRATION NUMBER: 35,367
? REFERENCE/DOCKET NUMBER: 015280-297100US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 576-0200
? TELEFAX: (415) 576-0300
? TELEX: <unknown>
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 8065 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: -
? LOCATION: 1...8065
? OTHER INFORMATION: Clone 22
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 116...1036
? OTHER INFORMATION: Clone 22 coding region
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 452...505
? OTHER INFORMATION: alternatively spliced portion
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 5595...5685
? OTHER INFORMATION: amplified region for genotyping
? SEQUENCE DESCRIPTION: SEQ ID NO: 6:
? US-09-091-952A-6
?
Query Match          26.7%; Score 352.2; DB 4; Length 8065;
Best Local Similarity 73.1%; Pred. No. 7.2e-58;
Matches 482; Conservative 0; Mismatches 168; Indels 9; Gaps 2;

```

```

Db 461 CTGTGGCTTTCAGACAGCGCCGACCGCGGTGG-----CGCCTCGAGATCATGCAT 514
Oy 746 CCGCTGGGCGCCACCGACCGCTGGCGGTGCGCCCTTCCGCCAGCGGGAGCGTTCCAC 805
Db 515 GCCCGGGGTCGACAGAGGTTTACAGCGCGCTCTTCATCCAGAGGATCGCTTCAGC 574
Oy 806 CGCTTCCAGCCACCTATCCGTACCTCAGACAGATGCACTGCACCCACCACCATCTCG 865
Db 575 CGCTTCCAGCCACCTATCCGTATGTGACAGCAGATGATCTTCTCCCATCATCTCC 634
Oy 866 CTGTACAGCGGGGAGAGCCCGCCACCTACCAGGGCCCTGACCTCCAGCTTCGGGAC 925
Db 635 CTGTCCAGCGGTGAAGAGCCACCTCTTACCAGGGCGCTGACCTGACAGCTCCGGGAC 694
Oy 926 CCGAGAGAGAGCTGGAACCTGAACCGGAGTGGGTGGCGGACCCCGCAACAGAACATC 985
Db 695 CCGAAGAGAGAGTGAACCTGAACCGAGAGTCCGTGAGGGCCCGCCCAACCGAACATA 754
Oy 986 TTGCACAGTACCTGATGATAGTCCAGG---CTGGGCGCGCCCTGCGCCCGCAGAGT 1042
Db 755 TTTCAGCATGATTTAATAGACATTTGATATAGCGGGGTTCCATGCCACCCAGCAGC 814
Oy 1043 AACTGGGCAATCAGCGCCAGCTGTACGCAAGCGGGGCGCATGAGAGGGCGCGCGCC 1102
Db 815 AACTGGGCAATCAGTGAAGAGCCTGACAGAGTAAAGGAGATGAGAGGGCGCACCCCG 874
Oy 1103 ACCTACAGCGAGTCACTGCGCACTACCGGGGCTCTTCCAGCAACAGACAGAGAG 1161
Db 875 ACATACAGCGAGTGAATGGCGCACCGCGCGCTCTTCTTCATCACACGAGCAG 933

```

## RESULT 3

```

US-09-091-952A-8
; Sequence 8, Application US/09091952A
; Patent No. 6458532

```

## GENERAL INFORMATION:

```

APPLICANT: Detera-Wadleigh, Sevilla D.

```

```

Gershon, Elliott S.

```

```

Badner, Judith A.

```

```

Goldin, Lynn R.

```

```

Berretini, Wade H.

```

```

Yoshikawa, Takeo

```

```

Sanders, Alan R.

```

```

Esterling, Lisa E.

```

```

TITLE OF INVENTION: Chromosomal Markers and Diagnostic
Tests for Manic-Depressive Illness

```

```

NUMBER OF SEQUENCES: 197

```

```

CORRESPONDENCE ADDRESS:

```

```

ADDRESS: Townsend and Townsend and Crew LLP

```

```

CITY: San Francisco

```

```

STATE: CA

```

```

COUNTRY: USA

```

```

ZIP: 94111-3834

```

```

COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Diskette

```

```

OPERATING SYSTEM: DOS

```

```

SOFTWARE: FastSeq for Windows Version 2.0

```

```

CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/09/091,952A

```

```

FILING DATE: 19-Apr-1999

```

```

CLASSIFICATION: <Unknown>

```

```

PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 60/029, 278

```

```

FILING DATE: 28-Oct-1996

```

```

APPLICATION NUMBER: PCT/US97/19381

```

```

FILING DATE: 28-Oct-1997

```

```

ATTORNEY/AGENT INFORMATION:

```

```

NAME: Smith, Timothy L.

```

```

REGISTRATION NUMBER: 35,367

```

```

REFERENCE/DOCKET NUMBER: 015280-297100US

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 867 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..867
; OTHER INFORMATION: Clone 22 isoform 2 alternatively
; spliced coding region
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
; US-09-091-952A-8

```

```

Query Match 21.5%; Score 284.2; DB 4; Length 867;
Best Local Similarity 68.7%; Pred. No. 3e-45;
Matches 453; Conservative 0; Mismatches 143; Indels 63; Gaps 2;

```

```

Oy 506 CAGAGCATGAGATCAGGAGCTGAGTTTGTTCAGATCATCATCGTGTGTATG 565
Db 166 CCGGCACTCTTCAACTCGAGCTGAGTTCGCCCAATCATCATCATCGTGTGTGTC 225
Oy 566 ATGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 625
Db 226 ACGGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 285
Oy 626 TTTCATGACCGCGGACAGCGAGGGGAGAGAGAGATGCGTCTCGAAGAGATGC 685
Db 286 TTTCATGACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 329
Oy 686 CTGTGGCCCTTGAGAGACAGATGTCAAGCAAGCAAGCAAGCAAGCAAGCAAG 745
Db 330 -----GCGCGCATCATGCAT 345
Oy 746 CCGCTGGGCGCCACCGACCGCTGCGCGTCCCGCTTCCCGCAGCGGAGCGTTCCAC 805
Db 346 GCCCGCGGCTCAGGAGAGAGTTCACAGCGCGCTCTTCATCCAGAGGATCGCTTCAGC 405
Oy 806 CGCTTCCAGCGCCACCTATCCGTACCTGACAGAGATCGACACCCACCATCTCG 865
Db 406 CGCTTCCAGCGCCACCTATCCGTATGTGACAGACAGATGATCTTCCCGACCATCTCC 465
Oy 866 CTGTACAGCGGGAGAGAGCGCCACCGTACAGAGGGCCCTGACCTCCAGCTTCGGGAC 925
Db 466 CTGTCCAGCGGTGAAGAGCCACCTCTTACAGAGGGGCGCTGACCTGACCTCGGGAC 525
Oy 926 CCGAGAGCAGAGTGAACCTGAACCGGAGTGGGTGGCGGACCCCGCAACAGAACATC 985
Db 526 CCGAAGAGAGAGTGAACCTGAACCGAGAGTCCGTGAGGGGCGCCCAACCGAACATA 585
Oy 986 TTGCACAGTACCTGATGATAGTCCAGG---CTGGGCGCGCCCTGCGCCCGCAGAGT 1042
Db 586 TTGCACAGTATTTAATAGACATTTGATATAGCGGGGTTCCATGCCACCGACGAC 645
Oy 1043 AACTGGGCAATCAGCGCCAGCTGTACGCGAGCGGGCGCCATGAGAGGGCGCGCGCC 1102
Db 646 AACTGGGCAATCAGTGAAGAGCCTGACAGAGTGAAGGAGAGTGAAGGGGCGCACCCCG 705
Oy 1103 ACCTACAGCGAGTCACTGCGCACTACCGGGGCTCTTCCAGCAGCAGAGAGAG 1161
Db 706 ACATACAGCGAGTGAATGGCGCACCGACCGCGCTCTTCTTCATCACACGAGCAG 764

```

## RESULT 4

```

US-08-804-227C-7/c
; Sequence 7, Application US/08804227C
; Patent No. 5876991

```

```

GENERAL INFORMATION:

```

APPLICANT: Dehoff, Bradley S.  
APPLICANT: Kustoss, Stuart A.  
APPLICANT: Rostock, Paul R., Jr.  
APPLICANT: Sutton, Kimberly L.  
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THOMAS G. PLANT 1501  
STREET: LILLY CORPORATE CENTER  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: ASCII(DOS) Text only  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,227C  
FILING DATE: February 21, 1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas G.  
REGISTRATION NUMBER: 35,784  
REFERENCE/DOCKET NUMBER: X-8231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-2459  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4437 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 350..14002  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 14046..20036  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 20110..31284  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31329..36071  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 36155..41830  
US-08-804-227C-7

Query Match 5.18; Score 68; DB 2; Length 44377;  
Best Local Similarity 49.2%; Pred. No. 0.00018;  
Matches 234; Conservative 0; Mismatches 240; Indels 2; Gaps 2;

QY 5 CGCGCTCTCGAGGAAACCGATCTCTTGACTTGAATGAGAGGAGGCGCGGCG 64  
Db 19499 CTGGGTGCGGACGTCCTCCGCGACCCGCGCGCAGCGTCCCGCGCGTCCGCGCGGCG 19440  
QY 65 GCGCGCGCGCGCGAGCGCTCGGCTGGGAAAGCTAGCGCGCAGAGGCTAGCCCGCGCG 124  
Db 19439 GCGGCGAGGGGCGACGCGACGCTCTGTCCTCCGCGCGCGCGCGCGCGCGCT 19380  
QY 125 GCG 184  
Db 19379 GCGGCGCTCCCGCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19320  
QY 185 CCGCGCGCGCTCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 244  
Db 19319 GCG 19261  
QY 245 ECCCG 304

Db 19260 CCGGCTCGACAGAGTTCAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19201  
QY 305 CCG 364  
Db 19200 CCG 19141  
QY 365 GCG 424  
Db 19140 AGCGGAGCGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19082  
QY 425 ATGGGGGTCAACAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
Db 19081 AAGGCGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19026

RESULT 5  
US-08-804-198-1/c  
Sequence 1, Application US/08804198  
Patent No. 5945320  
GENERAL INFORMATION:  
APPLICANT: Burgett, Stanley G.  
APPLICANT: Kustoss, Stuart A.  
APPLICANT: Rao, Nagara R.  
APPLICANT: Richardson, Mark A.  
APPLICANT: Rostock, Paul R., Jr.  
TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PAUL R. CANTRELL 1138  
STREET: LILLY CORPORATE CENTER  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh 7.0  
SOFTWARE: Microsoft Word 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,198  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CANTRELL, PAUL R.  
REGISTRATION NUMBER: 36,470  
REFERENCE/DOCKET NUMBER: P9113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3885  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4437 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 350..14002  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 14046..20036  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 20110..31284  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31329..36071  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 36155..41830  
US-08-804-198-1



```

Db 2828 TGGCTGGGCCCCGGGGGCTGCGCGCCAGCCGCCAGGGGTGGGGGCTTGGGGGCG 2769
OY 100 TAGCGGACAGAGGCTCAGCCCGCGCGGACAGCGCGCCCGCTGCCAGCCATTTCGGA 159
Db 2768 GGGCGGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2709
OY 160 GCGCAGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 219
Db 2708 GGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2649
OY 220 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 278
Db 2648 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2589
OY 279 AGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 338
Db 2588 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2529
OY 339 TCGG-GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 397
Db 2528 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2469
OY 398 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 457
Db 2468 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2409
OY 458 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 517
Db 2408 GCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2349
OY 518 ATCAGCGAGC 527
Db 2348 CTCACGGCGC 2339

```

## RESULT 8

US-08-843-659-1/c

Sequence 1, Application US/08843659

Patent No. 6218103

GENERAL INFORMATION:

APPLICANT: Leopardi, Roastio

TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White &amp; Durkee

CITY: Houston

STATE: Texas

COUNTRY: United States

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/843,659

FILING DATE: Concurrently Herewith

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4257 base pairs

Type: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-843-659-1

Query Match 5.1%; Score 67.6; DB 4; Length 4257;  
Best Local Similarity 48.8%; Pred. No. 0.0016;  
Matches 239; Conservative 0; Mismatches 249; Indels 2; Gaps 2;

```

OY 40 TGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 99
Db 2828 TGGCTGGGCCCCGGGGGCTGCGCGCCAGCCGCCAGGGGTGGGGGCTTGGGGGCG 2769
OY 100 TAGCGGACAGAGGCTCAGCCCGCGCGGACAGCGCGCCCGCTGCCAGCCATTTCGGA 159
Db 2768 GGGCGGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2709
OY 160 GCGCAGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 219
Db 2708 GGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2649
OY 220 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 278
Db 2648 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2589
OY 279 AGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 338
Db 2588 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2529
OY 339 TCGG-GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 397
Db 2528 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2469
OY 398 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 457
Db 2468 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2409
OY 458 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 517
Db 2408 GCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2349
OY 518 ATCAGCGAGC 527
Db 2348 CTCACGGCGC 2339

```

## RESULT 9

US-08-458-568A-11

Sequence 11, Application US/08458568A

Patent No. 5821339

GENERAL INFORMATION:

APPLICANT: Schaffer, Priscilla A.

TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &amp; No. 5821339r1s

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/458,568A

FILING DATE: 02-JUNE-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/055,146



```

RESULT 11
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007 .00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551

```

OTHER INFORMATION: "a" bases at various positions throughout the sequence  
OTHER INFORMATION: represent a, t, c or g  
us-09-103-840A-2

Query Match 5.0%; Score 65.4; DB 4; Length 4403765;  
Best Local Similarity 46.1%; Pred. No. 0.001;  
Matches 219; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

QY 1 CGACCGGCTCGAGAGGAAACCGATCCTTGGACTTGAATGAGAGAGAGCGG 60  
DB 3936911 CGCACCCCTTCCGCGCGCCGCCCTTACCGGCGGCCCGCGGCGCGCGCGC 3936852  
QY 61 CGGCGGCGCGCGCGCGCGCGCGCTGCGGAACTAGCGGAGAGAGAGCGCGC 120  
DB 3936851 CGGTGATATGCGCGCTTGGCCCGCTCCCGCGGACCGCGCGCGCGCGC 3936792  
QY 121 GCGCGGAGCGCGCGCGCGCGCTGCGAGCCATTTCCGAGCGCACCGCGCGCGC 180  
DB 3936791 CGGCACCGCGCGCGCGCGCGCGCTGCGAGCTTGTCCGCGCGCGCGCGC 3936732  
181 ACGCCCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGC 240  
DB 3936731 TCGC 3936672  
QY 241 TGAAGC 300  
DB 3936671 CGC 3936612  
QY 301 CTGC 360  
DB 3936611 CGGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 3936552  
QY 361 CCGC 420  
DB 3936551 CCGC 3936492  
QY 421 CTGATGAGGAGTCAACAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 475  
DB 3936491 CGGACGTGATGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 3936437

## RESULT 12

US-08-845-998-7/c  
Sequence 7, Application US/08845998  
Patent No. 5879892

## GENERAL INFORMATION:

APPLICANT: Van Baren, Nicolas  
APPLICANT: Coulie, Pierre G.  
APPLICANT: De Smet, Charles  
APPLICANT: Lucas, Sophie  
APPLICANT: Boon, Thierry  
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/845.998  
FILING DATE:

## CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Van Amsterdam, John R.  
REGISTRATION NUMBER: 40,212  
REFERENCE/DOCKET NUMBER: L0461/7008

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)720-3500  
TELEFAX: (617)720-2441  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4524 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:

NAME/KEY: CDS  
LOCATION: 174..1433  
US-08-845-998-7

Query Match 4.8%; Score 64; DB 2; Length 4524;  
Best Local Similarity 49.5%; Pred. No. 0.00074;  
Matches 218; Conservative 0; Mismatches 220; Indels 2; Gaps 2;

QY 166 CGCGCGGCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 225  
DB 897 CGACCTGC 839  
QY 226 CGCGGTCGC 285  
DB 838 GCGACCGC 779  
QY 286 CGAGC 345  
DB 778 TCGCAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 719  
QY 346 GC-CGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 404  
DB 718 CGC 659  
QY 405 CGCGCTCCATGACCGCTTGTATGAGGAGTCAACAGACCGCGCGCGCGCGCGC 464  
DB 658 CGC 599  
QY 465 CCAATGTCCTGCGACGTGCAAGTCAACAGCTCTTGTTCAGAGCATGAGATCAGC 524  
DB 598 CGC 539  
QY 525 AGCTGAGTTGTTTCAGATCATCATGATGATGATGATGATGATGATGATGATG 584  
DB 538 GAGGAGATGTTGTCAGACATGCGCGGTTGAGCGCGCTGATGATGATGATGATG 479  
QY 585 TCAGTGCCTGCTGAGCCAC 604  
DB 478 TGGTGTGC 459

## RESULT 13

US-09-206-537-7/c  
Sequence 7, Application US/09206537  
Patent No. 6130052

## GENERAL INFORMATION:

APPLICANT: Van Baren, Nicolas  
APPLICANT: Coulie, Pierre G.  
APPLICANT: De Smet, Charles  
APPLICANT: Lucas, Sophie  
APPLICANT: Boon, Thierry  
TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02210

```

: Patent NO.56271019
: GENERAL INFORMATION:
: APPLICANT: Van Baren, Nicolas
: APPLICANT: Coulie, Pierre G.
: APPLICANT: De Smet, Charles
: APPLICANT: Lucas, Sophie
: APPLICANT: Boon, Thierry
: TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: US
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/430,854
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/845,998
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Van Amsterdam, John R.
: REGISTRATION NUMBER: 40,212
: REFERENCE/DOCKET NUMBER: L0461/7008
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)720-3500
: TELEFAX: (617)720-2441
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4524 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 174..1433
: US-09-430-854-7
:
: Query Match 4.88; Score 64; DB 4; Length 4524;
: Best Local Similarity 49.5%; Pred. No. 0.00074;
: Matches 218; Conservative 0; Mismatches 220; Indels 2; Gaps 2;
:
: 166 CCGGGGACACGCGGAGCGCCCGGGGGCTGCGAGGGAGAGCGGGGGGGGCGACGGAG 225
: Db 897 CCACTGCGCGGCGAGCCCGCGCGCTGCGCTGCGCGCGC-GTGGTGGCGCGCCGCC 839
:
: 226 CGCGGTCCCGGCACTGAGCCCGCGGCGCCCGGGAACTTGGCGGGCAGACCGGACCGG 285
: Db 838 GCCACACACCGCGGGGTGGCGGACGCCGGAGCGGCAATGTTCAATGCGCGCGCGCGG 779
:
: 286 CGAGCGGGGGCGCGCTTCCCGCGCGCGCGCTCTCGATGCGGGGGCCCCCAAGTCCGGG 345
: Db 778 TGCACAGAGTGCGCGAGCTGTGCAATATGCGGGGTGAGGGTGCAGGACCGCCAGGAGC 719
:
: 346 GC-CGGCGCGGAGCCCGCGCGCGCGCGCGCGCGAGACCGCCCGCGCGCGCGCGCGCG 404
: Db 718 CCGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 659
:
: 405 CGCGCTCCATCACCAGCTTGATGAGGGGTCAACAGCAGCAGCGCGCGCGCGCGCGGAGC 464
: Db 658 CGCGCGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599

```

MOLECULE TYPE: DNA (genomic)

Job time : 1067.57 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 13:40:19 ; Search time 112.606 Seconds  
(without alignments)  
8233.708 Million cell updates/sec

Title: US-09-934-249-1

Perfect score: 1321

Sequence: 1 cgaccgcgcttcgcgagcga.....ctgcgttagtgtaaaagcag 1321

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapept 1.0

Searched: 501302 seqs, 350932545 residues

Optical number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1321	100.0	1321	10	US-09-934-249-1
2	861	65.2	861	10	US-09-934-249-3
3	803.2	60.8	969	9	US-09-796-753-55
4	790.4	59.8	1066	9	US-10-098-841-71
5	616.4	46.7	878	10	US-09-934-249-12
6	589.2	44.6	1583	10	US-10-000-2564-32
7	493.6	37.4	693	10	US-09-934-249-14
8	466.2	35.3	1713	9	US-09-796-753-57
9	349	26.4	8093	10	US-09-934-249-16
10	341.8	25.9	475	10	US-09-934-249-15
11	182	13.8	368	10	US-09-783-590-3464
12	78	5.9	4020	9	US-09-796-679-5
13	75.4	5.7	2561	12	US-10-023-529-48
14	75.4	5.7	2561	12	US-10-023-523-48
15	75.4	5.7	2561	12	US-10-023-523-48
16	72.2	5.5	2307	9	US-09-893-519A-87
17	72.2	5.5	2561	9	US-09-976-740-48
18	72.2	5.5	2561	12	US-10-023-529-48
19	72.2	5.5	2561	12	US-10-023-523-48

C	20	68	5.1	884.1	9	US-09-954-531-1351	Sequence 1351, Ap
C	21	67.6	5.1	4257	9	US-09-825-288A-1	Sequence 1, Appl
C	22	67.4	5.1	43058	10	US-09-954-445-292	Sequence 292, App
C	23	67.4	5.1	43058	10	US-09-954-456-529	Sequence 529, App
C	24	67.4	5.1	43058	10	US-09-880-107-3950	Sequence 3950, Ap
C	25	67	5.1	12733	9	US-10-032-393-47	Sequence 47, Appl
C	26	67	5.1	12739	9	US-10-032-393-8	Sequence 8, Appl
C	27	66.6	5.0	1614	9	US-09-976-740-45	Sequence 45, Appl
C	28	66.6	5.0	1614	12	US-10-023-529-45	Sequence 45, Appl
C	29	66.6	5.0	1614	12	US-10-023-523-45	Sequence 45, Appl
C	30	66.6	5.0	12425	9	US-09-976-740-50	Sequence 50, Appl
C	31	66.6	5.0	12425	12	US-10-023-529-50	Sequence 50, Appl
C	32	66.6	5.0	12425	12	US-10-023-523-50	Sequence 50, Appl
C	33	66.4	5.0	42999	9	US-09-799-462A-17	Sequence 17, Appl
C	34	66.4	5.0	42999	9	US-10-125-767-17	Sequence 17, Appl
C	35	66.4	5.0	42999	9	US-09-836-911A-17	Sequence 17, Appl
C	36	66	5.0	152331	9	US-10-095-407-16	Sequence 16, Appl
C	37	65.8	5.0	42999	9	US-09-799-462A-17	Sequence 17, Appl
C	38	65.8	5.0	42999	9	US-10-125-767-17	Sequence 17, Appl
C	39	65.8	5.0	42999	9	US-09-836-911A-17	Sequence 17, Appl
C	40	64.4	4.9	1614	12	US-10-023-529-45	Sequence 45, Appl
C	41	64.4	4.9	1614	12	US-10-023-523-45	Sequence 45, Appl
C	42	64.4	4.9	12425	9	US-09-976-740-50	Sequence 50, Appl
C	43	64.4	4.9	12425	12	US-10-023-529-50	Sequence 50, Appl
C	44	64.4	4.9	12425	12	US-10-023-523-50	Sequence 50, Appl
C	45	64.4	4.9	12425	12	US-10-023-523-50	Sequence 50, Appl

## ALIGNMENTS

RESULT 1  
US-09-934-249-1  
; Sequence 1, Application US/09934249  
; Patent No. US20020115081A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Richard T.  
; APPLICANT: Landschultz, Katherine T.  
; APPLICANT: Turt, Thomas G.  
; APPLICANT: Thompson, John F.  
; APPLICANT: Kennedy, Scott P.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS  
; FILE REFERENCE: P0738/7001/ERP/KA  
; CURRENT APPLICATION NUMBER: US/09/934, 249  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/227, 159  
; PRIOR FILING DATE: 2000-08-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1321  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (413)...(1273)  
US-09-934-249-1

Query Match 100.0%; Score 1321; DB 10; Length 1321;  
Best Local Similarity 100.0%; Pred. No. 2.7e+242;  
Matches 1321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCGCGGCTCTCGAGCGAACCAGTCCTTGACTGAATGAGGAGGAGCGG 60  
DB 1 CGACCGCGGCTCTCGAGCGAACCAGTCCTTGACTGAATGAGGAGGAGCGG 60  
QY 61 CGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
DB 61 CGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
QY 121 GGGGCGAGCG 180  
DB 121 GGGGCGAGCG 180

D	122	GGCGGAGCGCGCGCCCGTGCACGCCATTTTCGAGCGCACCGCGGCGACTGCGG	180
OY	181	ACGCCCGCGGGCTCCGAGGGAGAGCCGGGGGGGCCACGGAGCGCGTCCCGCGAC	240
D	181	ACGCCCGCGGGCTCCGAGGGAGAGCCGGGGGGGCCACGGAGCGCGTCCCGCGAC	240
OY	241	TGAGCCCGCGGGCCCGCGGGAACCTTGCGGCGACCCGAGCCGGGAGCGCGGGCGGC	300
D	241	TGAGCCCGCGGGCCCGCGGGAACCTTGCGGCGACCCGAGCCGGGAGCGCGGGCGGC	300
OY	301	CTCCCGCGCGCGCGCTCCGTGCATGCGGGGGCCCAAGCTCCGGGGCCCGCGAGCCCC	360
D	301	CTCCCGCGCGCGCGCTCCGTGCATGCGGGGGCCCAAGCTCCGGGGCCCGCGAGCCCC	360
OY	361	CCCGGGCGCGCCCGCGAGCCCGCCCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCG	420
D	361	CCCGGGCGCGCCCGCGAGCCCGCCCGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCG	420
OY	421	CTTGATGGGGGTCAACAGCACCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	480
D	421	CTTGATGGGGGTCAACAGCACCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	480
OY	481	GTCGAACCTGAAGGCTCTTTGTTCAGAGCATGGAATCAAGAGCTGAGTTGTTC	540
D	481	GTCGAACCTGAAGGCTCTTTGTTCAGAGCATGGAATCAAGAGCTGAGTTGTTC	540
OY	541	GATCATCATCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG	600
D	541	GATCATCATCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG	600
OY	601	CCACTGCAAGCTGTGTGACAGCGTCTTTATACGCGGACACAGCGAGGGGAGAGAGA	660
D	601	CCACTGCAAGCTGTGTGACAGCGTCTTTATACGCGGACACAGCGAGGGGAGAGAGA	660
OY	661	AATGATCCCTGTCTCAGAGAGATGCTGTGCGCCCTCGAGAGAGCAACTGTACAGCAAG	720
D	661	AATGATCCCTGTCTCAGAGAGATGCTGTGCGCCCTCGAGAGAGCAACTGTGTACAGCAAG	720
OY	721	AATCCAGAGCGGAGGCTACAGCGCCCGCGCTCGGCCACCGACGCGCTGCGTGCAGCG	780
D	721	AATCCAGAGCGGAGGCTACAGCGCCCGCGCTCGGCCACCGACGCGCTGCGTGCAGCG	780
OY	781	CTTGCGCCAGCGGAGCGCTTCACCGCTTCAGGCCACCTATCCGTACTCTGACAGCA	840
D	781	CTTGCGCCAGCGGAGCGCTTCACCGCTTCAGGCCACCTATCCGTACTCTGACAGCA	840
OY	841	GATGACCTGCGACCCACCACTCTGCTGTAGAGCGGGAGAGAGCCCGCCACTTACAGAG	900
D	841	GATGACCTGCGACCCACCACTCTGCTGTGTAGAGCGGGAGAGAGCCCGCCACTTACAGAG	900
OY	901	CCCTGCAACCTCAGCTTCGAGGACCCGAGAGCGAGCTGGAACCTGGGAGTCCGT	960
D	901	CCCTGCAACCTCAGCTTCGAGGACCCGAGAGCGAGCTGGAACCTGGGAGTCCGT	960
OY	961	GCGCGGACCCCAAGAGAAACAATCTTGACAGTGACCTGATGATGATGATGATGATG	1020
D	961	GCGCGGACCCCAAGAGAAACAATCTTGACAGTGACCTGATGATGATGATGATGATG	1020
OY	1021	CGGCGCCCTGCCCCCGACGATTAATCGGGGATTAAGCCCACTGTCAAGCGCACCGCG	1080
D	1021	CGGCGCCCTGCCCCCGACGATTAATCGGGGATTAAGCCCACTGTCTAAGCGCACCGCG	1080
OY	1081	GCGCATGAGGGGGCGCGCGCCACTTAAGCGAGGTCAATCGGCACATACCGGGGCTCTC	1140
D	1081	GCGCATGAGGGGGCGCGCGCCACTTAAGCGAGGTCAATCGGCACATACCGGGGCTCTC	1140
OY	1141	CTTCCAGCAGCAGCAGAGCAGTGGGGCGCGCTCTCTGAGAGGGACCGGCTCACCA	1200
D	1141	CTTCCAGCAGCAGCAGAGCAGTGGGGCGCGCTCTCTGAGAGGGACCGGCTCACCA	1200
OY	1201	CACACACATCCCGCCCTAGAGAGCGAGCCATCTGGAGCAAGAGAGATTAACAGAA	1260
D	1201	CACACACATCCCGCCCTAGAGAGCGAGCCATCTGGAGCAAGAGAGATTAACAGAA	1260

[illegible]



QY	1114	GGCTATGGGCGACATACCCGGGGGTCTCTTCCAGACACGAGAGACAGATGGGGCCGCCCTC	1173
Db	602	GGTATATGGGCGACATACCCGGGGGTCTCTTCCAGACACGAGAGACAGATGGGGCCGCCCTC	661
QY	1174	CTTGCTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCAT	1233
Db	662	CTTGCTGTGAGGGGAGCCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCAT	721
QY	1234	CTGGAGCGAAAGAGAGATTAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGGCGG	1293
Db	722	CTGGAGCGAAAGAGAGATTAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGGCGG	781
QY	1294	GGCTGGGGCTGGCGTAGGTGAAAGGGAG	1321
Db	782	GGCTGGGGCTGGCGTAGGTGAAAGGGAG	809

## RESULT 4

Sequence 71, Application US/10098841  
Publication No. US20020197679A1

**GENERAL INFORMATION:**

```

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feijian
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhilwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Dimaane, Radjoe T.
TITLE OF INVENTION: No. US20020197679a1el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2
CURRENT APPLICATION NUMBER: US/10,098,841
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: PL_FL-genes Version 1.0
SEQ ID NO 71
LENGTH: 1066
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (154)..(867)
US-10-098-841-71

```

Query Match	59.8%	Score 790.4	DB 9	Length 1066
Best Local Similarly	99.28%	Pred. No. 1.2e-141		
Matches 794	0	Mismatches 6	Indels 0	Gaps 0

[illegible]

Dp	223	GCACAGTGTACAGGACAGGAATCCACAGAGCCGAGGTTACAGCCCGCTCGGCCACCG	761
Qy	702	GCACAGTGTACAGGACAGGAATCCACAGAGCCGAGGTTACAGCCCGCTCGGCCACCG	761
Dp	293	GCACAGTGTACAGGACAGGAATCCACAGAGCCGAGGTTACAGCCCGCTCGGCCACCG	352
Qy	762	ACCGCCTGGCCGTCCGCCCTTGCGCCAGGGAGGCGTTCCACCGCTTCAGGCCACT	821
Dp	353	ACCGCCTGGCCGTCCGCCCTTGCGCCAGGGAGGCGTTCCACCGCTTCAGGCCACT	412
Qy	822	ATCGCTACCTGCACAGCAATGCATCGCACCCACATCTCGCTGCAGAGGGAGG	881
Dp	413	ATCGCTACCTGCACAGCAATGCATCGCACCCACATCTCGCTGCAGAGGGAGG	472
Qy	882	AGCCCCACCCCTACACAGGGCCCTGCACCCCTCAGCTTGCGGAGCCCGACAGCACTGG	941
Dp	473	AGCCCCACCCCTACACAGGGCCCTGCACCCCTCAGCTTGCGGAGCCCGACAGCACTGG	532
Qy	942	AACTCGAACCGGGAGTCGGTCGCGGCGACACCCCAAAACGAACATCTTCAGCACTGCATCA	1001
Dp	533	AACTCGAACCGGGAGTCGGTCGCGGCGACACCCCAAAACGAACATCTTCAGCACTGCATCA	592
Qy	1002	TGATATAGTCCAGGCTGGGGGCGCCCTGCCGCCCGACAGTACTGGGCTATCAGCGCA	1061
Dp	593	TGATATAGTCCAGGCTGGGGGCGCCCTGCCGCCCGACAGTACTGGGCTATCAGCGCA	652
Qy	1062	CGTGTACGGCAGCGGCGGGCGCATGGAGGGGCGCGCCACCTTACAGCAGAGTATCG	1121
Dp	653	CGTGTACGGCAGCGGCGGGCGCATGGAGGGGCGCGCCACCTTACAGCAGAGTATCG	712
Qy	1122	GCACACTACCGGGTCTCTCTTCACAGACACAGAGACAGTGGGCGCGCTCTCTGGTGG	1181
Dp	713	GCACACTACCGGGTCTCTCTTCACAGACACAGAGACAGTGGGCGCGCTCTCTGGTGG	772
Qy	1182	AGGGAGCCCGGCTCCACACACACATCGCGCCCTAAGAGAGCGCAGCATCTGGAGCA	1241
Dp	773	AGGGAGCCCGGCTCCACACACACATCGCGCCCTAAGAGAGCGCAGCATCTGGAGCA	832
Qy	1242	AAGAGAAGATAAACAGAAAGACACCTCTAGGGTCCCCAGGGGGCGGGCTGGGG	1301
Dp	833	AAGAGAAGATAAACAGAAAGACACCTCTAGGGTCCCCAGGGGGCGGGCTGGGG	892
Qy	1302	CTGCGTAGTGTAAGAGCAG	1321
Dp	893	CTGCGTAGTGTAAGAGCAG	912

## RESULT 5

```

US-09-934-249-12
Sequence 12, Application US/099344249
Patent No. US20020115081A1
GENERAL INFORMATION:
APPLICANT: Lee, Richard T.
APPLICANT: Landschulz, Katherine T.
APPLICANT: Turi, Thomas G.
APPLICANT: Thompson, John P.
APPLICANT: Kennedy, Scott P.
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
FILE REFERENCE: P0738/7001/ERP/KA
CURRENT APPLICATION NUMBER: US/09/934, 249
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 60/227, 159
PRIOR FILING DATE: 2000-08-22
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 878
TYPE: DNA
ORGANISM: Mus Musculus
FEATURE:
NAME/KEY: CDS

```



LOCATION: (20)...(841)  
US-09-934-249-12

QY	416	CACCGCTTGTGGGGGTCACACACACGCGCGCGCGCGGAGGACCAATGTCTTC	475
Db	11	CATGCGCTGTATGGGGGGTCAACGGCACCGCGCGCGCGCGCGGAGCCAAATGTCTTC	70
QY	476	TGCACGTGCAACTGTCAAAACGCTCTTTGTTCACAGCAGCATGAGATCACGAGCTGAGATT	535
Db	71	TGCCTGTGCAACTGTCCAGCGCTCTTTGTTCCTCCAGCATGAGATCACGAGCTGAGATT	130
QY	536	GTTTCAGTATCATCATCATCGGTGGTGGATGATGATGTGTGTGTGTATCATCGTGGCTG	595
Db	131	GTTGCAAAATCGTGGTCAATCGTGGATGATGATGATGTGTGTGTGTATCATCGTGGCTG	190
QY	596	CTGAGCCATCTACAAAGCTGTCTGTGCACGGTCTTCATCAGCGGACAGCCAGGGGGGAG	655
Db	191	CTGAGCCATCTACAAAGCTGTGACGCCGCGTCTTCATCAGCGGACAGCGCCAGGGGAG	250
QY	656	AGAGAAATATGCTCTGTCTCTAGAAGATGCTTGTGGCCTCGGAGAGCACTGTGAGCG	715
Db	251	AGAGACATATGATGTCTCTGTGGAAGATGCTTGTGGCCTCGGAGAGATCGGTGTGAGG-	309
QY	716	AACGGAATCCAGAGCGCGAGTCTACGCGCCGCTCGGCGCCACGACAGCGCTGGCGG	775
Db	310	--TGGAAATGGCGGAGGACCAAGGTCTATGCGCGCGCTCGGCGCCACTGACGACTCGCTG	367
QY	776	CCGCGCTTGGCCCAAGGGGGAGCGCTTCACACGCGCTTCACAGCCACTATCCGTAAC	835
Db	368	CCCGCTTTCATCAGCGCG-----AGCGGATTCACCAACCACTTACCTTACCTG	418
QY	836	CACGAGATTCGACCTGGCCACCCACCATCTGCGCTGTCAAGAGGGAGAGGCCACACCTAC	895
Db	419	CACGAAATTTGCTCTGCGCACCCACCATCTACTGTGTGATGGGAGAGGCCCCACCTTAC	478
QY	896	CAGGGCCCTGTCAACCCCTTCAGCTTCGGGACCCCGAGCAGCAGCTGTGAATGAACGGGAG	955
Db	479	CAGGGCCCTGTCAACCTCTCCAGCTACGGGACCTTGAGCAAGCTGTGAGCTGAACGGGAG	538
QY	956	TTCGTTGGGCGCACCCCAACAGAACCATTTTGAAGATGACTGTGATGATAGTGCAGG	1011
Db	539	TCTTGTGGCGGACCCCTTAACCGGACCATTTTGAAGATGACTTTTAACAGACACCAAG	598
QY	1016	CTGCGGCGGCGCTTCCCGCCACAGATGAATCGGGATCAGCGCCACGTCAGTACGGCAGC	1072
Db	599	CTGCGGCGGCGCTTCTCCCGCCAGAGTAATCGGGGATCAGCGCCACGTCCTACAGAC	658
QY	1076	GGCGGGCGCATGGAGGGGGCGCGCCACCTACAGCGAGGTGATCGGGCACTACCCGGGG	1133
Db	659	GGTGGGGCGCATGGAGGGGGCGCGCCACCTACAGGGAGGTGATTTGGCCACTAACCTGTGC	718
QY	1136	TCTCTCTTTCAGGACAGAGAGAGAGTGGGGCGCCCTTCCTGTGTGAGGGGAGCCGGGCTC	1199
Db	719	TCTCTCTTTCAGGACAGCAAAGTAACGGGCGCATCTCCCTGTGTAAAGGGAGCCGGGCTC	778
QY	1196	CACCAACACACATCGCGGCCCTTAGAGAGCGCACCATCTGTGAGCAAAAGAGATATA	1255
Db	779	CATCACTCGCACATTTGGCCCACTGGA-----GAAACAAGGAAAGGAGAA	823
QY	1256	CAGAAAGCACACCTCTCTTAGAGGTCCCAAGGGGGCGG	1293
Db	824	CAGAAAGGTCAACCCCTCTTAGAGATGGGGGCGCGGGCG	861

```

RESULT 6
US-10-000-256A-32
; Sequence 32, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION:

```

ORGANISM: Homo sapiens  
US-10-000-256A-32

[illegible]

RESULT 7  
US-09-934-249-14/c  
Sequence 14, Application US/09934249  
Patent No. US20020115081A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Richard T.  
APPLICANT: Landschulz, Katherine T.  
APPLICANT: Turti, Thomas G.  
APPLICANT: Thompson, John F.  
APPLICANT: Kennedy, Scott P.

TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
FILE OF INVENTION: CARDIOVASCULAR CONDITIONS  
FILE REFERENCE: P0738/7001/ERP/KA  
CURRENT APPLICATION NUMBER: US/09/934,249  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 60/227,159  
PRIOR FILING DATE: 2000-08-22  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 14  
LENGTH: 693  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (639)...(639)  
OTHER INFORMATION: a, c, g, or t/u  
US-09-934-249-14

Query Match 37.4%; Score 493.6; DB 10; Length 693;  
Best Local Similarity 97.4%; Pred. No. 2.4e-85;  
Matches 533; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 776 CCGGCTTGGCCGAGCGAGCGCTTCACCGCTTCACCGCCACCTATCCGTAAGTCTGAG 835  
DB 693 CCGGCTTGGCCGAGCGAGCGCTTCACCGCTTCACCGCCACCTATCCGTAAGTCTGAG 634  
QY 836 CACGAGATGACCTGCCACCCACCATCTGCTCTCAGAGGGGAGGAGCCGACCTAC 895  
DB 633 CACGAGATGACG-CGCGCCGACCATCTGCTCTCAGAGGGG--AGGACCCGACCTAC 577  
QY 896 CAGGCGCCCTGACCTTCAGCTTCGAGACCCGAGCAGCAGCTGGAAGTGAACCGGGAG 955  
DB 576 CAGGCGCCCTGACCTTCAGCTTCGAGACCCGAGCAGCAGCTGGAAGTGAACCGGGAG 517  
QY 956 TCGGTGCGCGACCCCAACAGACATCTGTCAGACAGTACCTGATGATGATGTCAG 1015  
DB 516 TCGGTGCGCGACCCCAACAGACATCTGTCAGACAGTACCTGATGATGATGTCAG 457  
QY 1016 CTGGGCGGCGCCCTGCGCCCGAGCATCTGCGGATCAGCGGCGCTACGCGAGC 1075  
DB 456 CTGGGCGGCGCCCTGCGCCCGAGCATCTGCGGATCAGCGGCGCTACGCGAGC 397  
QY 1076 GCGGCGCGATGAGAGGGGCGCGCCGACCTACAGCGAGTCACTGGGCGCTACCGGGG 1135  
DB 396 GCGGCGCGATGAGAGGGGCGCGCCGACCTACAGCGAGTCACTGGGCGCTACCGGGG 337  
QY 1136 TCTCTCTTCCACGACGAGAGGAGGCGCGCTCTCTGCTGGA-GGGGACCGGCT 1194  
DB 336 TCTCTCTTCCACGACGAGAGGAGGCGCGCTCTCTGCTGGA-GGGGACCGGCT 277  
QY 1195 CCACGACACACATGCGCCCTAGAGAGCGCAGCATCTGAGCAAGAGAGATAA 1254  
DB 276 CCGGCGACACACATGCGCCCTAGAGAGCGCAGCATCTGAGCAAGAGAGATAA 217  
QY 1255 ACAGAAAGACACCTCTCTAGAGTCCCGAGGGGGCGGGCTGAGTAA 1314  
DB 216 ACAGAAAGACACCTCTCTAGAGTCCCGAGGGGGCGGGCTGAGTAA 157  
QY 1315 AAGGAG 1321  
DB 156 AAGGAG 150

RESULT 8  
US-09-796-753-57  
Sequence 57, Application US/09796753  
Publication No. US20030027998A1  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
FILE REFERENCE: 7853-227-999  
CURRENT APPLICATION NUMBER: US/09/796,753

CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 09/183,175  
PRIOR FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: 09/223,094  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/223,546  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/224,246  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/259,388  
PRIOR FILING DATE: 1999-02-26  
PRIOR APPLICATION NUMBER: 60/122,458  
PRIOR FILING DATE: 1999-03-01  
PRIOR APPLICATION NUMBER: 09/312,359  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 09/336,536  
PRIOR FILING DATE: 1999-06-18  
PRIOR APPLICATION NUMBER: 09/342,687  
PRIOR FILING DATE: 1999-06-29  
PRIOR APPLICATION NUMBER: 09/345,464  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: 09/365,164  
PRIOR FILING DATE: 1999-07-30  
PRIOR APPLICATION NUMBER: 09/399,723  
PRIOR FILING DATE: 1999-09-20  
PRIOR APPLICATION NUMBER: 09/409,634  
PRIOR FILING DATE: 1999-09-30  
PRIOR APPLICATION NUMBER: 09/471,179  
PRIOR FILING DATE: 1999-12-23  
PRIOR APPLICATION NUMBER: 09/474,071  
PRIOR FILING DATE: 1999-12-29  
PRIOR APPLICATION NUMBER: 09/474,072  
PRIOR FILING DATE: 1999-12-29  
PRIOR APPLICATION NUMBER: 09/514,010  
PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: 09/516,745  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/572,002  
PRIOR FILING DATE: 2000-05-14  
PRIOR APPLICATION NUMBER: 09/597,993  
PRIOR FILING DATE: 2000-06-19  
PRIOR APPLICATION NUMBER: 09/599,596  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 09/630,334  
PRIOR FILING DATE: 2000-07-31  
PRIOR APPLICATION NUMBER: 09/606,565  
PRIOR FILING DATE: 2000-06-29  
PRIOR APPLICATION NUMBER: 09/606,317  
PRIOR FILING DATE: 2000-06-29  
PRIOR APPLICATION NUMBER: 09/665,666  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: 09/677,751  
PRIOR FILING DATE: 2000-09-30  
NUMBER OF SEQ ID NOS: 162  
SEQ ID NO 57  
LENGTH: 1713  
TYPE: DNA  
ORGANISM: Mouse  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (2)...(652)  
US-09-796-753-57

Query Match 35.3%; Score 466.2; DB 9; Length 1713;  
Best Local Similarity 83.9%; Pred. No. 4.1e-80;  
Matches 573; Conservative 0; Mismatches 83; Indels 27; Gaps 3;

QY 595 GGTGAGGCACTACAGCTGTCTCTGACGGGCGCTTCATCAGCGGACAGCGGAGGAG 654  
DB 1 GGTGAGGCACTACAGCTGTCTCTGACGGGCGCTTCATCAGCGGACAGCGGAGGAG 60  
QY 655 GAGGAGATGCGCTGCTCAGAGATGCTGTGCGGCTCGAGAGCAGACTGTCAG 714

Db 61 GAGAGAGATGACTGTCTCTGAGGAGGATGCTTGGCCCTCAGAGATACGATGTGACG 120  
QY 715 CAAGGAAATCCAGAGCGGAGGTCTACGCCCCGCTCGGCCCAACCGCCCTGGCCGT 774  
Db 121 ---TGGAAATCCGAGGACAGAGGTCTATGCCCGCTCGGCCCACTACCGCACTGCTGT 177  
QY 775 GCGGCCCCCTGAGCGGAGGAGCGCTTCCAGCCGCTTCCAGCCCACTATCCGTATGCTGCA 834  
Db 178 GCGGCCCCCTTATCCAGGGG-----AGCCGATTCCAAACCCCTTACCTTACCTGCA 228  
QY 835 GCAAGAAATGAGACCTGCGCACCCACCATCTGCTGTGACAGCGGAGAGAGCCCACTCA 894  
Db 229 GCAAGAAATGAGACCTGCGCACCCACCATCTGCTGTGATGAGGAGAGAGCCCACTCA 288  
QY 895 CCAGGCCCCCTGACCTCTCCAGCTTCCGAGCCCGGACAGCAGTGGAACTGAAACCGGGA 954  
Db 289 CCAGGCCCCCTGACCTCTCCAGCTTCCGAGCCCGGACAGCAGTGGAACTGAAACCGGGA 348  
QY 955 GTCGGTGCGGCGACACCCCAACAGAACATCTTGACAGTGTGATGATGATGCGCAG 1014  
b 349 ATCTGTGCGGCGACACCCCTTACCGGACCATCTTGTGACAGTACCTTATAGACACCAT 408  
QY 1015 GCTGGGCGGCCCCCTGCCCCCAGCAGTAACTGCGGATCAGCGCCACGCTTACGCGAG 1074  
Db 409 GCTGGGCGGCCCCCTGCCCCCAGCAGTAACTGCGGATCAGCGCCACGCTTACGAGCAG 468  
QY 1075 CGGGGGGCGATGAGAGGGGCGGCGCCCTTACAGGAGGTCTATGCGCCCTACCGGGG 1134  
Db 469 CGGGGGGCGATGAGAGGGGCGGCGCCCTTACAGGAGGTCTATGCGCCCTACCGGGG 528  
QY 1135 GTCTCTCTTCAGACAGCAGCAGCAGTGTGGGCGCCCTCTCTGTGAGAGGAGCGCGCT 1194  
Db 529 CTCTCTCTTCAGACAGCAGCAGCAGTGTGGGCGCCCTCTCTGTGAGAGGAGCGCGCT 588  
QY 1195 CCACACACACACATGCGCGCCCTTACAGAGCGACCATCTGAGCAAGAAAGATTA 1254  
Db 589 CCACTACCTGCGACATGCGCCCACTGAG-----GAACAAGAGAGAGAGAA 633  
QY 1255 ACAGAAAGGACACCTCTCTAGG 1277  
Db 634 ACAGAAAGGTACCCCTCTAGG 656

RESULT 9  
US-09-934-249-16  
; Sequence 16, Application US/09934249  
; Patent No. US20020115081A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Richard T.  
APPLICANT: Landschultz, Katherine T.  
APPLICANT: Turi, Thomas G.  
APPLICANT: Thompson, John F.  
APPLICANT: Kennedy, Scott P.  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
FILE REFERENCE: P0738/7001/ERP/KA  
CURRENT APPLICATION NUMBER: US/09/934,249  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 60/227,159  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 8093  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (6477)...(6477)  
OTHER INFORMATION: c or t/u  
NAME/KEY: unsure  
LOCATION: (6837)...(6837)  
OTHER INFORMATION: a or c

US-09-934-249-16  
Query Match 26.48; Score 349; DB 10; Length 8093;  
Best Local Similarity 73.78; Pred. No. 8.2e-58;  
Matches 474; Conservative 0; Mismatches 160; Indels 9; Gaps 2;

QY 522 CGGAGCTGGAATGTTGATGATCATCATCATGCTGTGATGATGATGATGATGATGATG 581  
Db 250 CGGAGCTGGAATGTTGATGATCATCATCATGCTGTGATGATGATGATGATGATGATG 309  
QY 582 TGATCAGTGTGCTCTGATGACCTACCAAGCTGTGACAGGCTCTTCAATCAGCCGACA 641  
Db 310 TCATCTGTGCTCTGATGACCTACCAAGCTGTGACAGGCTCTTCAATCAGCCGACA 369  
QY 642 GCCAGGGGCGGAGAGAGAGATGCTGTCTTCAAGAGATGCTGTGCTGTGCTGTGAG 701  
Db 370 ACCAGAGCCGAGGAGGAGAGAGAGGCTGTGCGAGAGAGGCTGTGCTGTGCTGTGAG 429  
QY 702 GCACAGTGTGAGGAGAGAGAGATCCAGAGCCGAGGCTTACGCCCCGCTGCGCCACCG 761  
Db 430 GCGGCGGACCGCGGCTGAG-----CGCTTGGAGATCATCATGCCCCGCGGCTCAGGG 483  
QY 762 ACCGCTGTGCGGCTGCGGCTTTCGCGGAGGAGGCGCTTCCACGCTTCCAGGCCACT 821  
Db 484 ACAGGCTTACAGCGCGGCTTTCATCCAGAGAGGAGATGCTTCCAGGCCACT 543  
QY 822 ATCGCTACTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 881  
Db 544 ACCCTTATGTGACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 603  
QY 882 AGCCCCACCTTACAGAGGCGGCTGACGCTTCCAGGCTTCCAGGAGCAGGAGCTG 941  
Db 604 AGCCACTCTCTTACAGAGGCGGCTGACGCTTCCAGGAGCAGGAGCTG 663  
QY 942 AACTGAACCGGAGATGCTGTGCGCACCCCAACAGAACCATCTTGCAGAGTACCTGA 1001  
Db 664 AACTGAACCGGAGATGCTGTGCGCACCCCAACAGAACCATCTTGCAGAGTACCTGA 723  
QY 1002 TGGATGTGCGAGG---CTGGGCGGCGGCTGCCCCCAGAGATGATGATGATGATG 1058  
Db 724 TAGACATTTGATATATAGCGGGGCTCATGCCCCCAGAGAGATGATGATGATGATG 783  
QY 1059 CCAGCTGCTAGGAGCGGCGGCGCATGAGAGGCGGCGGCTTCCAGCTTACAGGAGTCA 1118  
Db 784 CAAGCAGCTGAGAGATGAGAGGAGATGAGAGGCGGCGGCTTCCAGCTTACAGGAGTCA 843  
QY 1119 TGGGCACTACCGGGGCTCTCTTCCAGCAGCAGCAGAGGAG 1161  
Db 844 TGGGCACTACCGGCGGCTCTCTTCCAGCAGCAGCAGGAG 886

RESULT 10  
US-09-934-249-15  
; Sequence 15, Application US/09934249  
; Patent No. US20020115081A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Richard T.  
APPLICANT: Landschultz, Katherine T.  
APPLICANT: Turi, Thomas G.  
APPLICANT: Thompson, John F.  
APPLICANT: Kennedy, Scott P.  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
FILE REFERENCE: P0738/7001/ERP/KA  
CURRENT APPLICATION NUMBER: US/09/934,249  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 60/227,159  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 15  
LENGTH: 475  
TYPE: DNA

```

RESULT 11
US-09-783-590-3464
: Sequence 3464, Application US/09783590
: Patent No. US20020110850A1
:
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: FO-16, C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIORITY APPLICATION NUMBER: 08/420,856
PRIORITY FILING DATE: 1995-04-12
PRIORITY APPLICATION NUMBER: 08/346,731
PRIORITY FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3464
LENGTH: 368
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (103)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (225)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (279)

```

45 GAGGAGGAGCGCGCGCGCGCGCGGAGGCGCTCGGCTGGGGAAGCTAGCG 104





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 17, 2003, 11:34:54 ; Search time 2480.96 Seconds  
(without alignments)  
8623.366 Million cell updates/sec

Title: US-09-934-249-1  
Perfect score: 1321  
Sequence: 1 cgaccgcgcgtctgcgagcga.....ctgcgtagtgtaaaagcag 1321

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estda:  
2: em\_esthum:  
3: em\_estln:  
4: em\_estnu:  
5: em\_estov:  
6: em\_estpl:  
7: em\_estro:  
8: em\_hlc:  
9: gb\_est1:  
10: gb\_est2:  
11: gb\_hlc:  
12: gb\_est3:  
13: gb\_est4:  
14: gb\_est5:  
15: em\_estfun:  
16: em\_estom:  
17: gb\_gss:  
18: em\_gss\_hum:  
19: em\_gss\_inv:  
20: em\_gss\_pln:  
21: em\_gss\_vrt:  
22: em\_gss\_fun:  
23: em\_gss\_man:  
24: em\_gss\_mus:  
25: em\_gss\_other:  
26: em\_gss\_pro:  
27: em\_gss\_rnd:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	859.4	65.1	1007	9	AL578575
C 2	840.8	63.6	949	9	AL517150
C 3	766.8	58.0	967	14	BO641849
C 4	766	58.0	1046	14	BM922276
C 5	736.4	55.7	916	14	BO954555
C 6	726.6	55.0	805	9	AL558881

Result	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT	FEATURES
7	707	53.5	825	9	AL543170							
8	706.6	53.5	898	11	AL558882							
9	630.6	47.7	1207	11	AK008976							
10	618.6	46.8	1079	11	BC023092							
11	614.6	46.5	782	14	BQ015170							
12	607.4	46.0	609	14	BO636742							
13	578.4	43.8	890	14	BO690750							
14	567.4	43.0	729	14	BO575741							
15	564.8	42.8	730	14	BM677602							
16	550	41.6	551	13	BM141979							
17	529.8	40.1	894	13	BI851941							
18	501.6	38.0	588	13	BM483503							
19	493.6	37.4	693	9	AI761441							
20	493.4	37.4	655	14	BO691705							
21	486	36.8	1400	13	BM559329							
22	468.8	35.5	1280	14	BO681500							
23	468	35.4	857	12	BG333347							
24	468	35.4	974	10	BB624904							
25	461	34.9	763	13	BI646175							
26	448	33.9	626	14	BM974296							
27	444.2	33.6	990	14	BO691065							
28	441	33.4	461	14	BM712680							
29	437.8	33.1	633	14	BM714472							
30	430	32.6	841	13	BI156703							
31	428.2	32.4	655	13	BI853324							
32	422	31.9	844	14	BO686793							
33	417.6	31.6	641	10	AM071693							
34	417	31.6	559	12	BE855409							
35	416.2	31.5	587	13	BI083462							
36	411	31.1	582	10	BE553323							
37	407.4	30.8	674	9	AI972096							
38	406.8	30.8	570	14	BO575582							
39	406.2	30.7	619	9	AI742327							
40	403	30.5	467	10	BE66930							
41	401.6	30.4	900	12	BF036086							
42	392.2	29.7	990	12	BG675643							
43	392	29.7	464	12	BF026695							
44	388	29.4	436	13	BM482193							
45	379.6	28.7	588	9	AI377498							

## ALIGNMENTS

RESULT 1  
AL578575/c  
LOCUS  
DEFINITION  
AL578575 L7L\_NFL006.PL2 Homo sapiens cDNA clone CS00K001YC24 3  
prime, mRNA sequence.  
ACCESSION  
AL578575  
VERSION  
AL578575.1 GI:12942781  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1007)  
AUTHORS  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished (2001)  
CONTACT  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. 1007  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS00K001YC24"  
/clone\_lib="L7L\_NFL006.PL2"  
/tissue\_type="placenta"  
/note="Vector: PCWSPORT 6; Site\_1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>

BASE COUNT 160 a 293 c 344 g 200 t 10 others  
 ORIGIN

Query Match 65.1%; Score 859.4; DB 9; Length 1007;  
 Best Local Similarity 97.7%; Pred. No. 4.2e-144;  
 Matches 882; Conservative 9; Mismatches 10; Indels 2; Gaps 2;

```

OY 420 GCTTATGGGGGTCAACAGCACCGCGCGCGCGGGGAG-CCCAATCTCTCC 478
DB 1007 GCTTATGGGGGTCAACAGCACCGCGCGCGCGGGGAGCCCAATTTCTCTCC 948
479 AGGTCAACTGCAACGCTCTTTGTTCCAGAGCATGAGATCAGGAGCTGGAGTTGTT 538
947 AGGTCAACTGCAACGCTCTTTGTTCCAGAGCATGAGATCAGGAGCTGGAGTTGTT 888
OY 539 CAGATCATCATCATCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 598
DB 887 CAGATCATCATCATCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 828
599 AGCCACTCAAGCTCTCTGACGCTCTTCATCAGCCGCGACAGCAGGAGGCGAGAGA 658
DB 827 AGCCACMCACAGCTCTCTGACGCTCTTCATCAGCCGCGACAGCAGGAGGCGAGAGA 768
OY 659 CAGATGCTCTCTCTCTCAGAGAGATGCTGTCGAGAGAGCAGTGTACAGCAAC 718
DB 767 GAGATGCTCTCTCTCTCAGAGAGATGCTGTCGAGAGAGCAGTGTACAGCAAC 708
719 GGAATCCAGAGCGGAGAGTCTAGCGCCGCTGCGCCACCGACCGCTGCGCTGCGG 778
OY 707 GGAATCCAGAGCGGAGAGTCTAGCGCCGCTGCGCCACCGACCGCTGCGCTGCGG 649
779 CCGTTGCGGCGAGGAGGAGGCTTCCAGCGCTTCCAGCCACTATCCGACTGCGAGC 838
DB 648 CCGTTGCGGCGAGGAGGAGGCTTCCAGCGCTTCCAGCCACTATCCGACTGCGAGC 589
OY 839 GAGATGACCTGCTCCACCCACCATCTCTGCTGTCAGACGGGAGAGAGCCGCCCTACAG 898
DB 588 GAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 529
899 GAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 958
OY 528 AGCCCTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 469
959 GTGCGCGGACCCCAAGAGAACATCTTTCGACAGTGAAGTGAATGATGATGATGATG 1018
DB 468 GTGCGCGGACCCCAAGAGAACATCTTTCGACAGTGAAGTGAATGATGATGATG 409
OY 1019 GCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1078
DB 408 GCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 349
1079 GCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1138
DB 348 GCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 289
1139 TCTCTTCAGACACAGAGAGAGAGTGGCGCGCGCTCTCTCTCTCTCTCTCTCTCT 1198
DB 288 TCTCTTCAGACACAGAGAGAGAGTGGCGCGCGCTCTCTCTCTCTCTCTCTCT 229
OY 1199 CACACACACATCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1258
DB 228 MACACACACATCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 169
1259 AAAGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1318

```

DB 168 AAAGAACACCT 109  
 OY 1319 CAG 1321  
 DB 108 CAG 106

RESULT 2  
 LOCUS AL517150/c 949 bp mRNA linear EST 13-FEB-2001  
 DEFINITION AL517150 LTL\_NFL011.NBC1 Homo sapiens cDNA clone CS0DA008YB23 3  
 prime, mRNA sequence.  
 ACCESSION AL517150  
 VERSION AL517150.1 GI:12780643  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 949)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: [sequef@genoscope.cns.fr](mailto:sequef@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

FEATURES  
 Source  
 1. 949  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DA008YB23"  
 /clone\_1lb="LTL\_NFL011.NBC1"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>"

BASE COUNT 145 a 289 c 330 g 183 t 2 others  
 ORIGIN

Query Match 63.6%; Score 840.8; DB 9; Length 949;  
 Best Local Similarity 98.8%; Pred. No. 8.8e-141;  
 Matches 856; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```

OY 456 CCGGCGAGCCCAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 515
DB 949 CCGGCGAGCCCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 890
OY 516 AGATCAGGAGGTGAGTGTGTTGATCATCATCATCATCATCATCATCATCATCAT 575
DB 889 AGATCAGGAGGTGAGTGTGTTGATCATCATCATCATCATCATCATCATCATCAT 830
OY 576 TGTGTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 635
DB 829 TGTGTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 770
OY 636 GGCACAGCCAGGGCGGAGAGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 695
DB 769 GGCACAGCCAGGGCGGAGAGAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 710
OY 696 CGGAGAGCAGTGTCTCAGGACCAAGCAATCCAGAGCCGAGAGTCTACGCCCGCTG 755
DB 709 CGGAGAGCAGTGTCTCAGGACCAAGCAATCCAGAGCCGCA-GTCTACGCCCGCTG 651

```



OY	756	CCACCGACCGCTGGCCTGCGGCCCTTGCGCACAGGGAGGGCTTCCACGCCTTCGACG	815
Db	650	CCACCGACCGCTGGCCTGCGGCCCTTGCGCACAGGGAGGGCTTCCACGCCTTCGACG	591
OY	816	CCACCTATCCGTACCCTGCGACGACGAGATCGACTGCGCACCCACCATCTCGCTGTCCAGC	875
Db	590	CCACCTATCCGTACCCTGCGACGACGAGATCGACTGCGCACCCACCATCTCGCTGTCCAGC	531
OY	876	GGGAGGACCCCCCAACCTTACCAAGGGCCCCCTGACCCCTCCAGTTTGGGACCCCAGCAGC	935
Db	530	GGGAGGACCCCCCAACCTTACCAAGGGCCCCCTGACCCCTCCAGTTTGGGACCCCAGCAGC	471
OY	936	AGCTGGAACCTGAAACGGGGAGTCGGGGCGGACACCCCAAACAGAACCAATCTTCGACAGTG	995
Db	470	AGCTGGAACCTGAAACGGGGAGTCGGGGCGGACACCCCAAACAGAACCAATCTTCGACAGTG	411
OY	996	ACCTGATGATATAGTCCAGAGGCTGGGCGGCCCTTGCCCCCAGCAGTAAGTCTGGGACATCA	1055
Db	410	ACCTGATGATATAGTCCAGAGGCTGGGCGGCCCTTGCCCCCAGCAGTAAGTCTGGGACATCA	351
OY	1056	GCGCCACGTGCTACGCGCAGCGGGCGGCGCATGGAAGGGCGCGCCGACCTTACAGCGAAG	1115
Db	350	GCGCCACGTGCTACGCGCAGCGGGCGGCGCATGGAAGGGCGCGCCGACCTTACAGCGAAG	291
OY	1116	TCATGGGCGCATTACCGGGGGTCTCTCTTCGACGACACAGCAGTAGTGAGGCGCCCTCTCT	1175
Db	290	TCATGGGCGCATTACCGGGGGTCTCTCTTCGACGACACAGCAGTAGTGAGGCGCCCTCTCT	231
OY	1176	TGCTGAGAGGGGACCGCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCT	1235
Db	230	TGCTGAGAGGGGACCGCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCT	171
OY	1236	GGAGCAAGAAGAGATTAACAAGAAAGACACCTCTCTAGGGTCCCAAGGGGGCGCGG	1295
Db	170	GGAGCAAGAAGAGATTAACAAGAAAGACACCTCTCTAGGGTCCCAAGGGGGCGCGG	111
OY	1296	CTGGGGCTGCGTAGGTGAAAAAGCGAG	1321
Db	110	CTGGGGCTGCGTAGGTGAAAAAGCGAG	85

RESULT 3  
B0641849 967 bp mRNA linear EST 15-JUL-2002

LOCUS AGENCOURT\_8287174 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:6292265

DEFINITION 5', mRNA sequence.

ACCESSION B0641849

VERSION B0641849.1 GI:21766021

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
TITLE NIH-mgc http://mgc.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.  
Email: cgabds-remail.nhl.gov  
Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:  
<http://image.lmnl.gov>  
Plate: LCM2493 row: g column: 18  
High quality sequence stop: 571.  
Location/Qualifiers  
1..967  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

[illegible]

BM922276 1046 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOURT 6707077 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5754437  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM922276  
 VERSION BM922276.1 GI:19372655  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1046)  
 NIH-MGC http://mgc.ncl.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Plate: L1AM12791 row=n column: 06  
 High quality sequence stop: 671.  
 Location/Qualifiers  
 1..1046  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5754437"  
 /clone\_1ib="NIH\_MGC\_115"  
 /lab\_host="DH10B"  
 /note="Organ: Pooled brain, lung, testis. Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 021. Note: this is a NIH\_MGC Library."  
 BASE COUNT 205 a 391 c 298 g 150 t 2 others  
 ORIGIN  
 Query Match 58.0%; Score 766; DB 14; Length 1046;  
 Best Local Similarity 95.8%; Pred. No. 2,1e-127;  
 Matches 818; Conservative 0; Mismatches 32; Indels 4; Gaps 3;

707 GTGTACAGCAACGAATATCCAGAGCCGAGGTCTACGCCCGGCTCGGCCACGACCCG 766  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 GTGTACAGCAACGAATATCCAGAGCCGAGGTCTACGCCCGGCTCGGCCACGACCCG 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 767 CTGGCCGTGGCGCCCTTGTGCGCCAGCGGAGCGGCTTCCACCGCTTCCAGCCACTTATCCG 826  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 CTGGCCGTGGCGCCCTTGTGCGCCAGCGGAGCGGCTTCCACCGCTTCCAGCCACTTATCCG 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 827 TACCTGACGACGAGATGACCTTCCACCAATCTCTGCTGACAGCGGAGAGGCC 886  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 481 TACCTGACGACGAGATGACCTTCCACCAATCTCTGCTGACAGCGGAGAGGCC 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 887 CCACCTTACGAGGCGCCCTGACACCTTCACTTGGGAGCCCGAGCAGAGTGAATG 946  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 541 CCACCTTACGAGGCGCCCTGACACCTTCACTTGGGAGCCCGAGCAGAGTGAATG 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 947 AACCGGAGTGGTGGCGGCGACCCCAACAGAACATCTTTCGACATGACCTGATGAT 1006  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 601 AACCGGAGTGGTGGCGGCGACCCCAACAGAACATCTTTCGACATGACCTGATGAT 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1007 AGTGCAGGCTGGGCGGCGCCCTGCGCCCGCCAGCACTGAGGATGAGCGCCACTGTC 1066  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 AGTGCAGGCTGGGCGGCGCCCTGCGCCCGCCAGCACTGAGGATGAGCGCGACTGTC 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1067 TACGCGACGCGCGGCGGCGCATGAGGAGGCGCGCG-CCCACTACAGGAGTCTGCGCC 1124  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 721 TACGCGACGCGGCGGCGCGCATGAGGAGGCGCGCGCCCACTACAGGAGTCTGCGCC 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1125 ACTA--CCCGGGGCTCTCTTCCAGCAGCAGCAGAGAGTGGCGGCGCTCTGCTGA 1182  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 781 ACTAACCAGGGGGGCTCTCTTCCAGCAGCAGCAGATACCACTGAGGAGCGCGCTCTTGC 840  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1183 GGGGACCGCGCTCC 1196  
 || || || ||  
 Db 841 CTGGAAGGGAAC 854  
 || || || ||  
 RESULT 5 916 bp mRNA linear EST 21-AUG-2002  
 B0954555  
 LOCUS AGENCOURT 8825282 Lupski\_sciatic\_nerve Homo sapiens cDNA clone  
 DEFINITION IMAGE:6204609 5', mRNA sequence.  
 ACCESSION B0954555  
 VERSION B0954555.1 GI:22370033  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 916)  
 NIH-MGC http://mgc.ncl.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Plate: L1AM13626 row: c column: 10  
 High quality sequence stop: 669.  
 Location/Qualifiers  
 1..916  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6204609"  
 /clone\_1ib="Lupski\_sciatic\_nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"



```

|||||
Db 445 CCGGAGTCGGTGGCGGACCCCAACAGACCATTTTGACAGTACATGATGATG 386
|||
Qy 1009 TGGCAGGCTGGGCGGCCCCCTCCCGGAGCATGACTCGGGCATCGGCCACGTGCTA 1068
|||||
Db 385 TGCCAGGCTGGGCGGCCCCCTCCCGGAGCATGACTCGGGCATCGGCCACGTGCTA 326
|||||
Qy 1069 CGGCAGCGCGGGCGGCATGAGAGGGGCGCGCCGACCTACAGGAGGTCAATCGGCCACTA 1128
|||||
Db 325 CGGCAGCGCGGGCGGCATGAGAGGGGCGCGCCGACCTACAGGAGGTCAATCGGCCACTA 266
|||||
Qy 1129 CCGGGGGTCTCTCTTCAGCAGCAGCAGAGCATGAGGCGGCGCTCTTCTGTGAGGGGAC 1188
|||||
Db 265 CCGGGGGTCTCTCTTCAGCAGCAGCAGAGCATGAGGCGGCGCTCTTCTGTGAGGGGAC 206
|||||
Qy 1189 CCGGCTCCACACACACATCGCGCCCTTAGAGAGCGGCACCATCTGGAGCAAGAGAA 1248
|||||
Db 205 CCGGCTCCACACACACATCGCGCCCTTAGAGAGCGGCACCATCTGGAGCAAGAGAA 146
|||||
Db 1249 GGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGCGTGGGCTGCGTA 1308
|||||
Db 145 GGATTAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGCGTGGGCTGCGTA 86
|||||
Qy 1309 GGTGAAGAGCAG 1321
|||||
Db 85 GGTGAAGAGCAG 73

```

```

RESULT 7
AL543170 825 bp mRNA linear EST 16-FEB-2001
LOCUS AL543170 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1002Y103 5
DEFINITION prime, mRNA sequence.
ACCESSION AL543170
VERSION AL543170.1 GI:12875648
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 825)
L.I.W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
ATTRES Location/Qualifiers
SOURCE
1. 825
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1002Y103"
/clone_1bp="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Lite technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

```

```

BASE COUNT 125 a 327 c 229 g 128 t 16 others
ORIGIN

```

```

Query Match 53.5%; Score 707; DB 9; Length 825;
Best Local Similarity 92.6%; Pred. No. 7.3e-117;
Matches 760; Conservative 13; Mismatches 44; Indels 4; Gaps 3;
Qy 110 GGCTCAGCCCGGCGGAGCGCGCCCGCTGCGCAGCCCATTTTCGAGCCGACCCGC 169

```

```

|||||
Db 1 GGCTCAGCCCGGCGGAGCGCGCCCGCTGCGCAGCCCATTTTCGAGCCGACCCGC 60
|||||
Qy 170 GGGCACTGCGGAGCGCCCGGCGGCTGCCGAGGAGCGCCGGGGGGGGCGAGCGGAGGCG 229
|||||
Db 61 GGGCACTGCGGAGCGCCCGGCGGCTGCCGAGGAGCGCCGGGGGGGGCGAGCGGAGGCG 120
|||||
Qy 230 GTCCCGGCTGAGAGCGCGGGGGCGCGGGAACTTGGCGGCGAGCGCGGAG 289
|||||
Db 121 GTCCCGGCTGAGAGCGCGGGGGCGCGGGAACTTGGCGGCGAGCGCGGAG 180
|||||
Qy 290 CCGGGGCGGCTCCCGCGCGCGCGCC -TCTGATCGGGGGCCCAAGCTCCGGGGCC 348
|||||
Db 181 CCGGGGCGGCTCCCGCGCGCGCGCC -TCTGATCGGGGGCCCAAGCTCCGGGGCC 240
|||||
Qy 349 GCGCGAGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408
|||||
Db 241 GCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
|||||
Qy 409 GTCCATGACACCGCTTATGAGGGGTCAGAGCAGCGCGCGCGCGCGCGCGCGCG 468
|||||
Db 301 TTCCATGACACCGCTTATGAGGGGTCAGAGCAGCGCGCGCGCGCGCGCGCGCG 360
|||||
Qy 469 TGTCTCTGAC--GTGCAACTGCAAGCGCTTTTGTTCAGAGCATGAGATGAGATC 526
|||||
Db 361 TTCTCTCTGACCTTGTATGAGGGGTCAGAGCAGCGCGCGCGCGCGCGCGCGCG 420
|||||
Qy 527 CTGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 586
|||||
Db 421 CTGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
|||||
Qy 587 ACGTCTGTGAGCGCACTCAAGCTGTGTGACAGCGCTTCTATGAGCGCGCGCGCG 646
|||||
Db 481 ACGTCTGTGAGCGCACTCAAGCTGTGTGACAGCGCTTCTATGAGCGCGCGCGCG 540
|||||
Qy 647 GGGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 706
|||||
Db 541 GGGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
|||||
Qy 707 GTGTCAGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 766
|||||
Db 601 GTGTCAGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
|||||
Qy 767 CTGCGCTGCGCGCGCTTGTGCGCGAGGAGGCTTCCAGCGCTTCCAGCGCTTCCG 826
|||||
Db 661 CTGCGCTGCGCGCGCTTGTGCGCGAGGAGGCTTCCAGCGCTTCCAGCGCTTCCG 720
|||||
Qy 827 TACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 886
|||||
Db 721 TACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779
|||||
Qy 887 CCACCTACAGAGGCGCGCTGCAACCTTCCAGCTTGGGGAGCC 927
|||||
Db 780 CCACCTACAGAGGCGCGCTGCAACCTTCCAGCTTGGGGAGCC 820

```

```

RESULT 8
AL558882 898 bp mRNA linear EST 16-FEB-2001
LOCUS AL558882 LTI_NFL008.TC2 Homo sapiens cDNA clone CS0J015YF12 5
DEFINITION prime, mRNA sequence.
ACCESSION AL558882
VERSION AL558882.1 GI:12903838
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 898)
L.I.W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
unpublished (2001)
CONTACT: Genoscope

```

Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [segreff@genoscope.cns.fr](mailto:segreff@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

**FEATURES**  
**source**

```
1. .898
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="CS0DJ015YF12"
   /clone_lib="LTI_NFL008_TC2"
   /sex="male"
```

Rockville, Maryland 20850, USA; Fax: (1) 301 610 8371  
Email: [filang@lifetech.com](mailto:filang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT	196 a	310 c	260 g	126 t	6 others
ORIGIN					

Query Match	53.5%;	Score 706.6;	DB 9;	Length 898;
Best Local Similarity	99.3%;	Pred. No. 8.6e-117;		
Matches 728; Conservative	2;	Mismatches 1;	Indels 2;	Gaps 2

OY	589	GTGCTCTGTAGCCACTACAAAGTGTCGAGCGTCTTCATACAGCCGACACAGCAGG	648
Db	1	GTGCTCTGTAGCCACTACAAAGTGTCGAGCGTCTTCATACAGCCGACACAGCAGG	60
OY	649	GCAGAGAGAGAAAGATCCCTGTCTTCAGAGGATCCGTGGCCCTCGAGAGCAGT	708
Db	61	GCAGAGAGAGAAAGATCCCTGTCTTCAGAGGATCCGTGGCCCTCGAGAGCAGT	120
OY	709	GTACGAGAACGGAAATCCACAGCGCAGGTCTACGCGCCGCTCGGCCACACGACGGCT	768
Db	121	GTACGAGAACGGAAATCCACAGCGCAGGTCTACGCGCCGCTCGGCCACACGACGGCT	179
OY	769	GGCCGTCGCCGCTTTCGCCAGCGGGAGCGCTTTCACACGCTTTCACGCCACCTATTCGTA	828
Db	180	GGCCGTCGCCGCTTTCGCCAGCGGGAGCGCTTTCACACGCTTTCACGCCACCTATTCGTA	239
OY	829	CCTGCACAGCAGATGCACCTGCCACCCACATCTGCTGTACAGACGGGAGAGCCCC	888
Db	240	CCTGCACAGCAGATGCACCTGCCACCCACATCTGCTGTACAGACGGGAGAGCCCC	299
7	889	ACCTTACAGAGGCCCTCGACACCTCCAGCTTTCGGAGCCCGAGCAGCAGCTGGAATGTA	948
Db	300	ACCTTACAGAGGCCCTCGACACCTCCAGCTTTCGGAGCCCGAGCAGCAGCTGGAATGTA	359
OY	949	CCGGAGTCGGTGCGCACCCCCCAACAGAACCATCTTCGACAGTGACCTAGTGGATAG	1008
Db	360	CCGGAGTCGGTGCGCACCCCCCAACAGAACCATCTTCGACAGTGACCTAGTGGATAG	419
OY	1009	TGCGAGCTGGGGGGCCCCCGCCCCACACTTAATCTGGGGATACAGGCCACGTGCTA	1068
Db	420	TGCGAGCTGGGGGGCCCCCGCCCCACACTTAATCTGGGGATACAGGCCACGTGCTA	479
OY	1069	CGCGAGCGGGGGGCGATGTGAGAGGGGGCGCGCCACCTACAGCGAGGTATCGGCCTA	1128
Db	480	CGCGAGCGGGGGGCGATGTGAGAGGGGGCGCGCCACCTACAGCGAGGTATCGGCCTA	539
OY	1129	CCCGGGGTCCTCTTCCAGCACACGACGAGCAGTGGGCGGCTCTTCTGTGAGGGGAC	1188
Db	540	CCCGGGGTCCTCTTCCAGCACACGACGAGCAGTGGGCGGCTCTTCTGTGAGGGGAC	599
OY	1189	CCGGTCACACACACACATCTGGGGCCCTTAGAGAGGCGACGCAATGTGGAGCAAGAGTA	1248
Db	600	CCGGTCACACACACACATCTGGGGCCCTTAGAGAGGCGA - SCATCTGAGCAAAAGATA	658
OY	1249	GGATTAACAGAAAGACACCTCTCTTAGGCTTCCACAGGGGGGCGGGCTGGGGCTGCGTA	1308

Accession	Sequence	Length
Db 659	GGATTAACAGAAAGACACCTCTCTAGGGTCCCAAGGGGGCCGGGCTGGGGCTGGCTA	718
Qy 1309	GCTGAAAAGGCAG	1321
Db 719	GGTCAAAAAGGCAG	731

RESULT 9  
AK008976

LOCUS	AK008976	1207 bp	mRNA	linear	HTC 19-JAN-2002
DEFINITION	Mus musculus adult male stomach cDNA, RIKEN full-length enriched				

ACCESSION	AK008976
VERSION	AK008976.1
KEYWORDS	GI:12843488
SOURCE	HTC; CAP trapper.
	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA

ORGANISM

REFERENCE

1

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus

AUTHORS

**TITLE** Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to

## AUTHORS

Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., and Tashiro, H. 2003. The structure of the *Arabidopsis thaliana* *AtHKT1* protein in the plasma membrane. *Plant Cell Physiol.* 44:1033-1041.

TITLE RIKEN integrated sequence analysis (RISA) system--384  
sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

## REFERENCE

## AUTHORS

TITLE  
 JOURNAL  
 MEDLINE  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (5821), 685-690 (2001)  
 21085660

PUBMED  
REFERENCE  
AUTHORS

11217851  
5 (bases 1 to 1207)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Butt,C.,  
Carinini,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanaoka,T.,  
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiroka,T., Horii,F.,  
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koys,S.,  
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishii,K., Nomura,K.,  
Numataki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,  
Salto,H., Salto,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,  
Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Tanahashi,F.,  
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,  
Yasunishi,A., Yoshida,K., Yoshino,M., Kuratsutsu,M. and  
Hayashizaki,Y.

TITLE  
JOURNAL

Direct Submission  
Physiol Cell Mol Biol (2000) Yoshida Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGGAGAAGATCCCAAGACCTCTTTTCTTTTTTTTAA 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 5.0 and subtraction Robt  
= 25.0. Second strand cDNA was prepared with the primer adapter of  
sequence [5'  
GAGAGAGATTTCGACGTATTATAATTATTAATCCCCCCCCCCC 3']. cDNA was cleaved  
with XhoI and SacI. Cloning sites, 5' end: XhoI; 3' end: SacI.  
Host: SOLR.

FEATURES  
SOURCE

Location/Qualifiers  
1..1207  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="RANTOM,DB:2210418102"  
/db\_xref="MGD:MGI:1902457"  
/db\_xref="taxon:10090"  
/clone="2210418102"  
/sex="male"  
/tissue\_type="stomach"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
391..1173  
/note="Nedd4 WW binding protein 4  
data source:MGD, source key:MGI:1929600, evidence:ISS  
putative"

CDS

BASE COUNT

Origin

Query Match 47.7%; Score 630.6; DB 11; Length 1207;  
Best Local Similarity 82.4%; Pred. No. 3.3e-103;  
Matches 782; Conservative 0; Mismatches 139; Indels 28; Gaps 4;

[illegible]



ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarste, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 36 Row: h Column: 6  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein  
This clone has the following problem: frame shifted.  
Location/Qualifiers  
1..1079  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="CZECH II"  
/clone="IMAGE:3989996"  
/tissue\_type="Mammary tumor metastatized to lung. Tumor  
arose spontaneously from a senescent normal mammary  
(clonal) outgrowth infected with the virus MMTV."  
/clone\_1id="NCI\_CGAP\_Lu29"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 271 a 341 c 289 g 178 t  
ORIGIN  
Query Match 46.8%; Score 618.6; DB 11; Length 1079;  
Best Local Similarity 84.5%; Pred. No. 4.6e-101;  
Matches 741; Conservative 0; Mismatches 109; Indels 27; Gaps 3;  
QY 417 ACCGCTGATGGGGTCAACAGCACCGCGCGCGCGGCGAGCCCAATGTCTCT 476  
|||||  
Db 15 ACCCTTGATGGGGTCAACAGCACCGCGCGCGCGGCGAGCCCAATGTCTCT 74  
|||||  
QY 477 GCAGGTGAATGGAACGCTCTGTTCAGAGCATGAGATCAGGAGTGTGAGTTG 536  
|||||  
Db 75 GCGGTGCAATGCGCAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 134  
|||||  
QY 537 TTCGATCATCATGCTGTGTTGATGATGATGATGATGATGATGATGATGATGAT 596  
|||||  
Db 135 TCGAATTCGTGTATGTGTATGATGATGATGATGATGATGATGATGATGATGAT 194  
|||||  
QY 597 TGACCCACTACAGCTGTCTCAGCGTCTTCATCAGCCGCGACAGCGGCGGAGA 656  
|||||  
Db 195 TGACCCACTACAGCTGTCTCAGCGTCTTCATCAGCCGCGACAGCGGCGGAGA 254  
|||||  
QY 657 GAGAAGATGCGCTGTCGCAAGAGATGCGTGTGCGCTCGAGAGACATGTGAGCA 716  
|||||  
Db 255 GAGACGATGAGCTGTCTCGAAGAGATGCGTGTGCGCTCGAGAGATGAGTGTGAG 312  
|||||  
QY 717 ACGGATCCAGAGCCGAGGTCTACGCGCGCTCGGCGCCAGCAGCGCTGCGCTGTC 776  
|||||

Db 313 -TGAATGCCGAGACACAGAGTATATGCCCGCCCTCGGCCCATCTAGCCAGTCTGCTG 371  
QY 777 CGCCCTTCCGCCAGCGGAGGCTTCCACCGCTTCCAGCCACCTATCCGATCTGAGC 836  
|||||  
Db 372 CCCCCTTATCCAGCG-----ACCGATTCCAAACCACTTACCTTACCTGAGC 422  
|||||  
QY 837 ACGATCGACCTCCACCCACCATCTCGCTGAGAGGAGAGAGAGAGAGAGAGAGAG 896  
|||||  
Db 423 ACGAATATGCCCTGACACCCACCATCTCTGATGTGATGAGGAGAGAGAGAGAGAG 482  
|||||  
QY 897 AGGCGCCCTGACCTCCAGCTTGGGAGCCCGAGCAGCAGCTGAACTGAACGGGAGT 956  
|||||  
Db 483 AGGCGCCCTGACCTCCAGCTTGGGAGCCCGAGCAGCAGCTGAACTGAACGGGAGT 542  
|||||  
QY 957 CGGTGCGGCGACCCCAACAGACATCTTGCAGAGAGAGAGAGAGAGAGAGAGAG 1016  
|||||  
Db 543 CTGTGCGGCGACCCCTTACCGAGCATCTTGCAGAGAGAGAGAGAGAGAGAGAGAG 602  
|||||  
QY 1017 TGGCGCGCCCTGCGCCCGCCAGCACTAAGTACGAGCAGCGCCAGCTGAGCGAGC 1076  
|||||  
Db 603 TGGGCGGCGCCCTGTCTCCCGCAGAGTACTGCGGATCAGGCGCAGCTGACAGCAG 662  
|||||  
QY 1077 GCGGCGCGCATGAGAGGCGCGCGCCACCTACAGCAGAGTATGCGCATACCGGGGT 1136  
|||||  
Db 663 GTGGCGCGATGAGAGGCGCGCGCCCGCCACCTACAGCAGAGTATGCGCATACCG 722  
|||||  
QY 1137 CCTCTTCCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196  
|||||  
Db 723 CCTCTTCCAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782  
|||||  
QY 1197 ACCACACACATCGCGCGCCCTTACAGAGCAGCATGTGCGAAGAGAGATTAAC 1256  
|||||  
Db 783 ATCACTGCGACATGCGCCACTGCA-----GAACAAGAGAGAGAGAGAAC 827  
|||||  
QY 1257 AGAAGGACACCTCTGTAGGGTCCCGCAGGGGGCGC 1293  
|||||  
Db 828 AGAAGGTCACCCCTCTAGAGAGTGGGGCGGGGGC 864  
|||||  
RESULT 11  
B0015170/c 782 bp mRNA linear EST 26-MAR-2002  
LOCUS  
DEFINITION  
UI-H-EDI-axv-k-20-0-UI-s1 NCI\_CGAP\_EDI Homo sapiens cDNA clone  
IMAGE:5834635 3', mRNA sequence.  
ACCESSION  
B0015170  
VERSION  
B0015170.1 GI:19740071  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 782)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Seq primer: M13 FORWARD  
POLYA=Yes.  
FEATURES  
source  
1..782  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5834635"  
/clone\_1id="NCI\_CGAP\_EDI"  
/tissue\_type="Chondrosarcoma"

		/dev_stage="Adult"	
		/lab_host="DH10B (Life Technologies)"	
		(Pharmacia) with a modified polylinker. Site 1: EcoR I; Site 2: Not I; NCI-CGAP-ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996.	
		First strand cDNA synthesis was primed with an oligo-dT ligated to an EcoR I adaptor. Double stranded cDNA was cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (GT)18 tail. The sequence tag for this library is GCTCAAGGCT.	
		TAG LIB=01-H-ED1	
		TAG TISSUE=chondrosarcoma	
		TAG_SEQ=CGTCAAGGCT	
BASE COUNT	109 a 223 c 271 g 176 t 3 others		
Query Match	46.5%; Score 614.6; DB 14; Length 782;		
Best Local Similarity	98.9%; Pred. No. 2.4e-100;		
Matches 617; Conservative	0; Mismatches 7; Indels 0; Gaps 0;		
QY 698	GAGAGCAGCTGCAGGACGAGGATCCAGAGCGAGTCTACGCCCGCTCGGCC 757		
DB 782	GAGAGCAGCTGCAGGACGAGGATCCAGAGCGAGTCTACGCCCGCTCGGCC 723		
QY 758	ACCGACCGCTGCGCGCTGCGCGCTGCGCGAGGGAGCGCTTCACACCGCTTCACACCG 817		
DB 722	ACCGACCGCTGCGCGCTGCGCGCTGCGCGAGGGAGCGCTTCACACCGCTTCACACCG 663		
QY 818	ACCTATCGTACCTGACAGAGATGACCTGCGACACCATCTGCTGACAGAGCG 877		
DB 662	ACCTATCGTACCTGACAGAGATGACCTGCGACACCATCTGCTGACAGAGCG 603		
QY 878	GAGAGCGCCCGACCTTACAGGCGCCCTGACACCTCCAGCTTGGGAGACCGCGACGAG 937		
DB 602	GAGAGCGCCCGACCTTACAGGCGCCCTGACACCTCCAGCTTGGGAGACCGCGACGAG 543		
QY 938	CTGGAACGAGCGGGAGCTGGTGGCGGACCCCGCAACAGAACATCTTGCACAGTGAC 997		
DB 542	CTGGAACGAGCGGGAGCTGGTGGCGGACCCCGCAACAGAACATCTTGCACAGTGAC 483		
QY 998	CTGATGATAGTGCAGAGCTGGCGGCGCCCTGCGCCCGCAGAGTAACTCGGGCAATCAGC 1057		
DB 482	CTGATGATAGTGCAGAGCTGGCGGCGCCCTGCGCCCGCAGAGTAACTCGGGCAATCAGC 423		
QY 1058	GCCACGTCATACGAGCGCGCGCCCATGAGAGGGCGCCCGCCACCTACAGCAGGTC 1117		
DB 422	GCCACGTCATACGAGCGCGCGCCCATGAGAGGGCGCCCGCCACCTACAGCAGGTC 363		
QY 1118	ATCGGCGCACTACCGCGGGGTCTCTTCCAGCAGCAGCAGAGAGTGGCGCGCTCTCTTG 1177		
DB 362	ATCGGCGCACTACCGCGGGGTCTCTTCCAGCAGCAGCAGAGAGTGGCGCGCTCTCTTG 303		
QY 1178	CTGGAGGGAGACCGCGCTCCACACACACATCGCGCCCTAGAGAGGCGAGCATCTGG 1237		
DB 302	CTGGAGGGAGACCGCGCTCCACACACACATCGCGCCCTAGAGAGGCGAGCATCTGG 243		
QY 1238	AGCAAGAGAGAGATTAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGGCGGGGCT 1297		
DB 242	AGCAAGAGAGAGATTAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGGCGGGGCT 183		
QY 1298	GGGGCTCGGTAGGTAAAGAGCAG 1321		
DB 182	GGGGCTCGGTAGGTAAAGAGCAG 159		

---

LOCUS	B0636742	609 bp	mRNA	linear	EST 15-JUL-2002
DEFINITION	hd13h06.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he				
ACCESSION	B0636742				
VERSION	B0636742.1				
KEYWORDS	EST				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Wistow G., Bernstein S.L., Wyatt M.K., Ray S., Behal A., Touchman J.W., Bouffard G., Smith D. and Peterson K.				
TITLE	Expressed sequence tag analysis of human retina for the NEIRBank project. Retbindin, an abundant, novel retinal cDNA and alternative splicing of other retina-preferred gene transcripts				
JOURNAL	Mol. Vis. 8 (4), (2002) In press				
COMMENT	Contact: Wistow G Section on Molecular Structure and Function National Eye Institute 6/331, NIH, Bethesda, MD 20892-2740, USA Tel: 301 402 3452 Fax: 301 496 0078 Email: gwaemehellix.nih.gov Plate: 13 row: h column: 06 Seq primer: M13R1 reverse primer (ABI).				
FEATURES	source				
	1..609				
	Location/Qualifiers				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="hd13h06"				
	/clone_lib="Human Retina cDNA (Un-normalized, unamplified)"				
	/hd/he				
	/tissue_type="Retina"				
	/dev_stage="Adult"				
	/lab_host="EMD10B"				
	/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue was dissected from two 80 year old donors with no observed eye disease. 100ug of total RNA was used for library construction. A directionally cloned cDNA library in the pSPORT1 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the Superscript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor (5'-pGACTGATCTGATCGGCGCGCGCCGCTT)15-3' J. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."				
BASE COUNT	114 a 238 c 182 g 75 t				
ORIGIN					
Query Match	46.0%; Score 607.4; DB 14; Length 609;				
Best Local Similarity	99.8%; Pred. No. 4.7e-99;				
Matches 608; Conservative	0; Mismatches 1; Indels 0; Gaps 0;				
QY 631	CAGCGGCAAGCAGGCGGAGAGAGATCCCTGTCCTCAGAGATGCTCTGTG 690				
DB 1	CAGCGGCAAGCAGGCGGAGAGAGATGCTCTGTCCTCAGAGATGCTCTGTG 60				
QY 691	GGCTTGGAGACAGATGTCAGGCAACGCAATCCAGAGCGCGAGGCTTACGCGCGCC 750				
DB 61	GGCTTGGAGACAGATGTCAGGCAACGCAATCCAGAGCGCGAGGCTTACGCGCGCC 120				
QY 751	TCGGCCACACGCGCTGCGCGCTGCGCGCTTGGCCACGCGGAGCGCTTCAACGCGCTT 810				
DB 121	TCGGCCACACGCGCTGCGCGCTGCGCGCTTGGCCACGCGGAGCGCTTCAACGCGCTT 180				
QY 811	CCAGCCACCTATCCGTAACGTCAGACAGAGATCGACCTGCCACCACTCTCGGTGTC 870				
DB 181	CCAGCCACCTATCCGTAACGTCAGACAGAGATCGACCTGCCACCACTCTCGGTGTC 240				
QY 871	AGAGGGGAGAGGAGCCCGACCGCTTACGAGGGCGCTGACCGCTTACGCTTGGGAGCCCGCA 930				

RESULT 12  
B0636742





DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 Seq primer: M13 FORWARD  
 POLYA=Yes

# FEATURES

Source

Location/Qualifiers

1. 729

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-H-E21-bdg-h-14-0-UI"

/clone\_1lb="NCI.CGAP.Ch2"

/tissue\_type="Chondrosarcoma Grade II"

/dev\_stage="Adult"

/lab\_host="DH10B (Life Technologies)"

/note="Organ: Left Pelvis; Vector: pRT73-Pac (Pharmacia)

with a modified polylinker. Site\_1: EcoR I; Site\_2: Not I;

NCI.CGAP.Ch2 is a normalized cDNA library containing the

following tissue(s): Chondrosarcoma Grade II. The library

was constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pRT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

TGATCAGCT.

TAG\_L1b-UI-H-E21

TAG\_TISSUE=grade-2-chondrosarcoma

TAG\_SEQ=ATCTATATAG"

BASE COUNT 101 a 213 c 253 g 162 t

ORIGIN

Query Match 43.0%; Score 567.4; DB 14; Length 729;

Best Local Similarity 99.8%; Pred. No. 6.6e-92;

Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

753 GGGCCACGACCGCTGGCGCTGGCGCCCTGGCGCGGAGAGCGCTTCACCGCTTCC 812

Db 729 GGGCCACGACCGCTGGCGCTGGCGCCCTGGCGCGGAGAGCGCTTCACCGCTTCC 670

Qy 813 AGCCACCTATCCGTACTCTGACGACGAGATCGACCTGCACCCACCATCTCGCTGAG 872

Db 669 AGCCACCTATCCGTACTCTGACGACGAGATCGACCTGCACCCACCATCTCGCTGAG 610

Qy 873 ACGGGGAGAGCCGCCACCTACGAGGCGCCCTGACAGCTTCGCGAGCCCGAGC 932

609 ACGGGGAGAGCCGCCACCTACGAGGCGCCCTGACAGCTTCGCGAGCCCGAGC 550

Qy 933 AGCAGCTGGAAGTGAACCGGGAGTGGTGGCGACCCCAAGAACATCTCGAGCA 992

Db 549 AGCAGCTGGAAGTGAACCGGGAGTGGTGGCGACCCCAAGAACATCTCGAGCA 490

Qy 993 GTGACCTGATGATGTGCGAGGCGGGGCGCCCTGCGCCCGACGAGTAAGTGGGCA 1052

Db 489 GTGACCTGATGATGTGCGAGGCGGGGCGCCCTGCGCCCGACGAGTAAGTGGGCA 430

Qy 1053 TCAGCCCGACAGTGTGAGCGACGGCGCGCATGAGAGGGGCGCGCCCGACGAGCGG 1112

Db 429 TCAGCCCGACAGTGTGAGCGACGGCGCGCATGAGAGGGGCGCGCCCGACGAGCGG 370

Qy 1113 AGGTATCGGACCTATACCGGGGCTCTCTTCAGACAGCAGAGAGAGTGGGCGGCT 1172

Db 369 AGGTATCGGACCTATACCGGGGCTCTCTTCAGACAGCAGAGAGAGTGGGCGGCT 310

Qy 1173 CCTTCCTGAGAGGAGCCGGCTCCACACACATCGGCGCCCTAGAGAGCGGACCA 1232

Db 309 CCTTCCTGAGAGGAGCCGGCTCCACACACATCGGCGCCCTAGAGAGCGGACCA 250

Qy 1233 TCTGGAGCAAGAGATGAAGATGAAGAGACACCTCTCTAGAGGCTCCCGAGGGGGCC 1292

Db 249 TCTGGAGCAAGAGATGAAGATGAAGAGACACCTCTCTAGAGGCTCCCGAGGGGGCC 190

Qy 1293 GGGCTGGGGCTGCGTAGTGTGAAGAGCGAC 1321

Db 189 GGGCTGGGGCTGCGTAGTGTGAAGAGCGAC 161

## RESULT 15

BM677602/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

BM677602 730 bp mRNA Linear EST 27-FEB-2002

UI-E-E01-ald-p-09-0-UI-s1 UI-E-E01 Homo sapiens cDNA clone

UI-E-E01-ald-p-09-0-UI 3', mRNA sequence.

BM677602.1 GI:18987498

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 730)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: Two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9555

Email: msosares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1. 730

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UI-E-E01-ald-p-09-0-UI"

/clone\_1lb="UI-E-E01"

/tissue\_type="fetal eye"

/dev\_stage="fetal"

/lab\_host="DH10B (Life Technologies) (TI phage resistant)"

/note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a

modified polylinker. Site\_1: EcoR I; Site\_2: Not I;

UI-E-E01 is a normalized cDNA library containing the

following tissue(s): fetal eye. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA,

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pRT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

CGGCTATAGC. This library was created for the program, Gene

Discovery in the Visual System, supported by National Eye

Institute (NEI).

TAG\_L1b-UI-E-E01

TAG\_TISSUE=human fetal eye

TAG\_SEQ=CGGCTATAGC"

BASE COUNT 103 a 210 c 251 g 164 t 2 others

ORIGIN

Query Match 42.8%; Score 564.8; DB 14; Length 730;

Best Local Similarity 99.3%; Pred. No. 1.9e-91;

Matches 566; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
OY 752 CGGCCACGACCGCTGGCCGCTTCCGCCAGCGGAGGCGCTTCCACCGCTTC 811
Db 730 CGGCCACGACCGCTTGGCCGCTTCCGCCAGCGGAGGCGCTTCCACCGCTTC 671
OY 812 CAGCCACCTATCCGTACCTGCAGCAGATGACCTGCACCCACCATCTCGCTGCA 871
Db 670 CAGCCACCTATCCGTACCTGCAGCAGATGACCTGCACCCACCATCTCGCTGCA 611
OY 872 GACGGGGAGAGCCCGCACCTACCAGGGCCCTGCACCTTCAGCTTGGGAGCCCGAG 931
Db 610 GACGGGGAGAGCCCGCACCTACCAGGGCCCTGCACCTTCAGCTTGGGAGCCCGAG 551
OY 932 CAGCAGCTGAGTGAACCGGAGTGGGCGGCGACCCCAAGAACCATCTTCGAC 991
Db 550 CAGCAGCTGAGTGAACCGGAGTGGGCGGCGACCCCAAGAACCATCTTCGAC 491
OY 992 AGTGACCTGATGATAGTGCAGAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 1051
Db 490 AGTGACCTGATGATAGTGCAGAGGCTGGGCGGCCCTGCCCCCAGCAGTAACTCGGGC 431
OY 1052 ATCAGCGCCACGTGCTACGGCAGCGGCGGCGCATGAGGGGCGCGCCCACTACAGC 1111
Db 430 ATCAGCGCCACGTGCTACGGCAGCGGCGGCGCATGAGGGGCGCGCCCACTACAGC 371
OY 1112 GAGGTATCGGCGCACTACCGGGGCTCTTCCAGCACACGAGAGAGTGGGCGGCC 1171
Db 370 GAGGTATCGGCGCACTACCGGGGCTCTTCCAGCACACGAGAGAGTGGGCGGCC 311
OY 1172 TCCTTGCTGAGAGGGACCGGCTCCACCAACACATCGGCGCCCTAGAGAGCGCAGCC 1231
Db 310 TCCTTGCTGAGAGGGACCGGCTCCACCAACACATCGGCGCCCTAGAGAGCGCAGCC 251
OY 1232 ATCTGAGACAAAGAGAGGATTAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGGC 1291
Db 250 ATCTGAGACAAAGAGAGGATTAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGGC 191
OY 1292 CGGGCTGGGGCTGGCTAGGTGTAAGGCGAG 1321
Db 190 CGGGCTGGGGCTGGCTAGGTGTAAGGCGAG 161
```

Search completed: March 17, 2003, 13:40:12  
Job time : 2492.96 secs

**THIS PAGE BLANK (USPTO)**